

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-345785

(43)Date of publication of application : 05.12.2003

(51)Int.Cl.

G06F 17/21

G06F 17/60

// G09B 19/00

(21)Application number : 2002-151821

(71)Applicant : FUKUDA TOYOSHI

SEEDWIN:KK

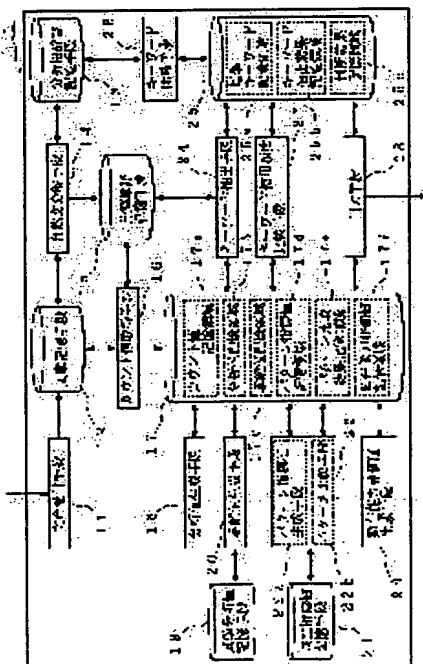
(22)Date of filing : 27.05.2002

(72)Inventor : FUKUDA TOYOSHI

MANTOKU CHIKAKO

NAGASHIMA HIROKO

(54) SYSTEM AND PROGRAM FOR EVALUATING ABILITY



(57)Abstract:

PROBLEM TO BE SOLVED: To objectively evaluate an ability of a writer based on the contents of a text.

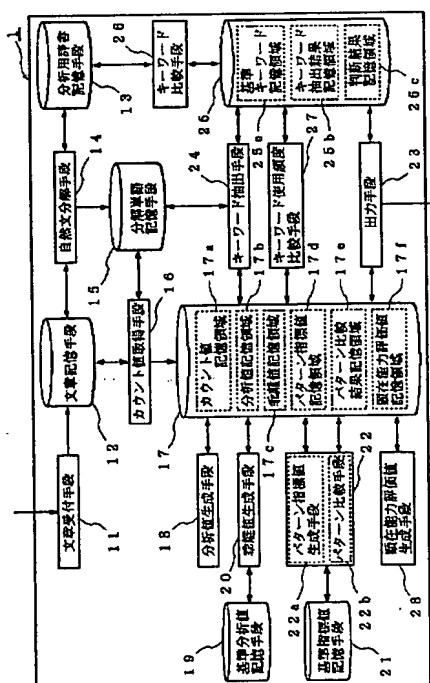
SOLUTION: An ability evaluating system includes: a text receiving means 11 for receiving the text; an analysis dictionary storage means 13 for stipulating reference unit words to be reference when a natural sentence is divided into words; a natural sentence dividing means 14 for dividing the natural sentence into the plurality of words; a division word storage means 15 for storing the plurality of words obtained by dividing the natural sentence; a count value obtaining means 16 for obtaining a number concerning the words as a count value about the text; an analysis value generating means 18 for generating a plurality of analysis values based on the count value; a

divergence value generating means 20 for generating a plurality of divergence values based on a reference analysis value read from a reference analysis value storage means 19, and the analysis values; a

divergence value comparing means 22 for comparing relations between the plurality of divergence values based on a reference index value read from a reference index value storage means 21; and an output means 23 for outputting a comparison result obtained by the divergence value comparing means 22.

(11)特許出願公開番号
特開2003-345785
(P2003-345785A)

(51)Int.Cl. ⁷	識別記号	F I	テーマコード [*] (参考)
G 0 6 F 17/21	5 5 0	G 0 6 F 17/21	5 5 0 A 5 B 0 0 9
17/60	1 5 8	17/60	1 5 8
// G 0 9 B 19/00		G 0 9 B 19/00	H



【特許請求の範囲】

【請求項1】 自然文からなる文章を受け付ける文章受付手段と、

当該自然文を単語に分解する基準となる基準単位単語を規定した分析用辞書記憶手段と、

当該分析用辞書記憶手段を参照して、前記基準単位単語に相当する複数の単語に前記自然文を分解する自然文分解手段と、

当該自然文分解手段が自然文を分解して得られた複数の単語を記憶する分解単語記憶手段と、

当該分解単語記憶手段が記憶した複数の単語を参照し、少なくとも当該単語にかかわる数を前記文章についてのカウント値として取得するカウント値取得手段と、

当該カウント値取得手段が取得したカウント値に基づき複数の分析値を生成する分析値生成手段と、

当該分析値の基準分析値を記憶する基準分析値記憶手段と、

当該基準分析値記憶手段から読み出した基準分析値と、生成された前記分析値と、に基づき当該基準分析値に対する当該分析値の乖離の程度を示す複数の乖離値を生成する乖離値生成手段と、

当該複数の乖離値についての比較基準を記憶する基準指標値記憶手段と、

当該基準指標値記憶手段から読み出した基準指標値に基づき、当該乖離値生成手段が生成した複数の前記乖離値同士の相対関係を比較する乖離値比較手段と、

当該乖離値比較手段が行った比較結果を出力する出力手段と、を備えること特徴とする能力評価システム。

【請求項2】 請求項1に記載の能力評価システムにおいて、

前記乖離値比較手段は、

前記乖離値生成手段が生成した複数の乖離値に基づき、乖離値相互の相対関係を表すパターン指標値を生成するパターン指標値生成手段と、

前記基準指標値記憶手段から読み出した基準指標値と当該パターン指標値と、を比較するパターン比較手段と、を有し、

前記出力手段は、当該パターン比較手段が行った比較結果を出力することを特徴とする能力評価システム。

【請求項3】 前記パターン指標値生成手段は、前記相対関係として、一の乖離値を基準として他の乖離値との差を用い、当該差の値に基づき各乖離値に対応するパターン指標値を生成することを特徴とする請求項2に記載の能力評価システム。

【請求項4】 請求項1乃至請求項3に記載の能力評価システムにおいて、前記分析用辞書記憶手段は、少なくとも単語の同義語もしくは類義語のいずれかを規定する比較用辞書記憶領域を備えるとともに、

前記自然文分解手段が自然文を分解して得られた複数の単語の中から、キーワードを抽出するキーワード抽出手

段と、

当該キーワード抽出手段が抽出したキーワードを文章毎に比較し、前記比較用辞書記憶領域を参照して同義語または類義語の関係にあるか、もしくはキーワードが一致するか、の判断を行うキーワード比較手段と、備え、前記出力手段は、当該キーワード比較手段が行った当該判断の結果を出力することを特徴とする能力評価システム。

【請求項5】 請求項4に記載の能力評価システムにおいて、さらに、

同義語または類義語の関係にあるか、あるいは、一致すると前記キーワード比較手段が判断したキーワードについて、前記分析値に基づき文章毎の使用頻度を比較するキーワード使用頻度比較手段を備え、

前記出力手段は、当該キーワード使用頻度比較手段が得た判断結果を出力することを特徴とする能力評価システム。

【請求項6】 請求項1乃至請求項5に記載の能力評価システムにおいて、

前記乖離値生成手段が生成した複数の乖離値に基づき複数の顕在能力評価値を生成する顕在能力評価値生成手段を備え、

前記出力手段は、当該顕在能力評価値を出力することを特徴とする能力評価システム。

【請求項7】 情報の受付手段と、情報の記憶手段と、これら各手段の動作を制御する処理手段と、を備え、

当該記憶手段が、文章を構成する自然文を単語に分解する基準となる基準単位単語を規定した分析用辞書記憶手段と、当該自然文を分解して得られた複数の単語を記憶

する分解単語記憶手段と、当該文章に対応する分析値の基準分析値を記憶する基準分析値記憶手段と、当該分析値に基づき得られる乖離値についての比較基準を記憶する基準指標値記憶手段と、を有する能力評価システムにおいて、

前記処理手段に、

前記受付手段を介して自然文からなる文章を受け付けるステップと、

前記分析用辞書記憶手段を参照して、前記基準単位単語に相当する複数の単語に前記自然文を分解するとともに、自然文を分解して得られた複数の単語を前記分解単語記憶手段に記憶するステップと、

当該分解単語記憶手段が記憶した複数の単語を参照し、少なくとも当該単語にかかわる数をカウント値として取得するステップと、

当該カウント値に基づき複数の前記分析値を生成するステップと、

生成した分析値と前記基準分析値記憶手段が記憶する基準分析値とに基づき、当該基準分析値に対する乖離の程度を示す複数の乖離値を生成するステップと、

前記基準指標値記憶手段から読み出した基準指標値に基

づき、生成した複数の前記乖離値同士の相対関係を比較するステップと、

当該相対関係についての比較結果を出力するステップと、を実行させることを特徴とする能力評価プログラム。

【請求項8】 請求項7に記載のプログラムにおいて、前記処理手段に、前記乖離値同士の相対関係を比較するステップとして、複数の前記乖離値に基づき、乖離値相互の相対関係を表すパターン指標値を生成するステップと、

当該生成したパターン指標値と、前記基準指標値記憶手段から読み出した基準指標値とを比較するステップと、を実行させることを特徴とする能力評価プログラム。

【請求項9】 請求項8に記載の能力評価プログラムにおいて、

前記処理手段に、前記相対関係として、一の乖離値を基準として他の乖離値との差を算出し、当該差の値に基づき各乖離値に対応するパターン指標値を生成するステップを実行させることを特徴とする能力評価プログラム。

【請求項10】 請求項7乃至請求項9に記載の能力評価プログラムにおいて、前記分析用辞書記憶手段は、少なくとも単語の同義語もしくは類義語のいずれかを規定する比較用辞書記憶領域を備えるものであって、

前記処理手段に、

前記自然文を分解して得られた複数の単語の中から、キーワードを抽出するステップと、

当該キーワードを文章毎に比較し、前記比較用辞書記憶領域を参照して同義語または類義語の関係にあるか、もしくはキーワードが一致するか、の判断を行うステップと、

当該判断の結果を出力するステップと、を実行させることを特徴とする能力評価プログラム。

【請求項11】 請求項10に記載の能力評価プログラムにおいて、さらに、

同義語または類義語の関係にあるか、あるいは、一致すると前記キーワード比較手段が判断したキーワードについて、前記分析値に基づき文章毎の使用頻度を比較するステップと、

当該使用頻度の比較結果を出力するステップと、を実行させることを特徴とする能力評価プログラム。

【請求項12】 請求項7乃至請求項11に記載の能力評価プログラムにおいて、前記処理手段に、前記複数の乖離値に基づき複数の顕在能力評価値を生成するステップと、

当該顕在能力評価値を出力するステップと、を実行させることを特徴とする能力評価プログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、客観的な数値的評価を行う能力評価システムおよび能力評価プログラムに

係り、特に、文章を用いた能力評価システムおよび能力評価プログラムに関する。

【0002】

【従来の技術】従来から、特に企業における人の評価に際しては、特定のテーマに対して記述させた論文を個人の能力評価の材料とすることが多い。例えば、入社試験におけるエントリーシートの一項目として800文字以上などの条件を付けて文章を書かせたり、社内での昇格試験に際して、事業部における戦略などを問う小論文を課したりしている。

【0003】これら特定のテーマに対して記述させた文章は、人事担当者や上級管理職などにより評価される。そして、これら評価者による評価は、特開2002-41726号公報にあるように、電子情報として管理され、必要に応じて閲覧などが可能なシステムに反映されることになるのが一般的である。

【0004】

【発明が解決しようとする課題】しかしながら、上述した特開2002-41726号公報にあるような一般的な情報管理システムでは、たとえ、複数人で評価を行い個人の主観を極力排除したとしても、所詮、文章の評価は人が行うものであり、評価の客観性を実現することは不可能である。

【0005】加えて、文章による評価を行うという手法は、上述したような人事関連だけに限られず、研修能力や営業能力などの判断を行う場面でも十分に採用され得るものであり、結果、それぞれの能力に応じた評価基準を確立しなければならない。また、市場動向が急激に変化する昨今の経済情勢の下では、求められる能力が容易に変化するものであり、それにともない、能力評価のための評価基準を変更することが当然に求められる。

【0006】そこで、人の能力を客観的に評価できるとともに、能力の評価基準を柔軟に変更できる新しい能力評価システムおよび能力評価プログラムの実現が望まれている。

【0007】

【発明の目的】本発明は、かかる要請に鑑み、特定のテーマについて書かれた文章に基づき、書き手の能力を客観的に評価することができる能力評価システムおよび能力評価プログラムを提供することを目的とする。また、評価基準を柔軟に変更することにより多様な能力評価に対応できる能力評価システムおよび能力評価プログラムを提供することを目的とする。

【0008】

【課題を解決するための手段】上記目的を達成するため、請求項1に記載の発明では、自然文からなる文章を受け付ける文章受付手段と、当該自然文を単語に分解する基準となる基準単位単語を規定した分析用辞書記憶手段と、当該分析用辞書記憶手段を参照して、前記基準単位単語に相当する複数の単語に前記自然文を分解する自

然文分解手段と、当該自然文分解手段が自然文を分解して得られた複数の単語を記憶する分解単語記憶手段と、当該分解単語記憶手段が記憶した複数の単語を参照し、少なくとも当該単語にかかわる数を前記文章についてのカウント値として取得するカウント値取得手段と、当該カウント値取得手段が取得したカウント値に基づき複数の分析値を生成する分析値生成手段と、当該分析値の基準分析値を記憶する基準分析値記憶手段と、当該基準分析値記憶手段から読み出した基準分析値と、生成された前記分析値と、に基づき当該基準分析値に対する当該分析値の乖離の程度を示す複数の乖離値を生成する乖離値生成手段と、当該複数の乖離値についての比較基準を記憶する基準指標値記憶手段と、当該基準指標値記憶手段から読み出した基準指標値に基づき、当該乖離値生成手段が生成した複数の前記乖離値同士の相対関係を比較する乖離値比較手段と、当該乖離値比較手段が行った比較結果を出力する出力手段と、を備える構成を採用する。

【0009】本発明によると、分解された複数の単語に基づき得られる、所定の基準分析値に対する乖離値を用いて乖離値同士の相対関係を比較するので、客観的に文章内容の評価を行うことができる。ひいては、複数の文章について、それぞれ比較結果を出力することにより、文章同士の相対評価も容易に行うことができる。

【0010】また、請求項2に記載の発明では、請求項1に記載の能力評価システムにおいて、前記乖離値比較手段は、前記乖離値生成手段が生成した複数の乖離値に基づき、乖離値相互の相対関係を表すパターン指標値を生成するパターン指標値生成手段と、前記基準指標値記憶手段から読み出した基準指標値と当該パターン指標値と、を比較するパターン比較手段と、を有し、前記出力手段は、当該パターン比較手段が行った比較結果を出力する構成を採用する。

【0011】本発明によると、乖離値の相対関係を表すパターン指標値を生成するので、より文章の持つあいまいさを加味した上で、文章のバランスについての的確な比較を行うことができる。

【0012】また、請求項3に記載の発明では、前記パターン指標値生成手段は、前記相対関係として、一の乖離値を基準として他の乖離値との差を用い、当該差の値に基づき各乖離値に対応するパターン指標値を生成する構成を採用する。

【0013】本発明によると、一の乖離値を基準として相対関係をみるので、乖離値の相対関係がより明瞭となる普遍化ができ、また、比較パターン数を少なくすることにより、マッチング処理の高速化が図れる。

【0014】また、請求項4に記載の発明では、請求項1乃至請求項3に記載の能力評価システムにおいて、前記分析用辞書記憶手段は、少なくとも単語の同義語もしくは類義語のいずれかを規定する比較用辞書記憶領域を備えるとともに、前記自然文分解手段が自然文を分解し

て得られた複数の単語の中から、キーワードを抽出するキーワード抽出手段と、当該キーワード抽出手段が抽出したキーワードを文章毎に比較し、前記比較用辞書記憶領域を参照して同義語または類義語の関係にあるか、もしくはキーワードが一致するか、の判断を行うキーワード比較手段と、備え、前記出力手段は、当該キーワード比較手段が行った当該判断の結果を出力する構成を採用する。

【0015】本発明によると、分解された複数の単語から抽出されたキーワードの同義語や類義語まで比較可能なので、単語の意味内容まで加味した精度の高い文章内容の比較評価を行うことができる。

【0016】また、請求項5に記載の発明では、請求項4に記載の能力評価システムにおいて、さらに、同義語または類義語の関係にあるか、あるいは、一致すると前記キーワード比較手段が判断したキーワードについて、前記分析値に基づき文章毎の使用頻度を比較するキーワード使用頻度比較手段を備え、前記出力手段は、当該キーワード使用頻度比較手段が得た判断結果を出力する構成を採用する。

【0017】本発明によると、キーワードの文章中で使用された頻度に基づく比較を行うので、文章中における書き手の単語の位置づけまでを含めた評価を実現する。

【0018】また、請求項6に記載の発明では、請求項1乃至請求項5に記載の能力評価システムにおいて、前記乖離値生成手段が生成した複数の乖離値に基づき複数の顕在能力評価値を生成する顕在能力評価値生成手段を備え、前記出力手段は、当該顕在能力評価値を出力する構成を採用する。

【0019】本発明によると、乖離値に基づき顕在能力評価値を生成、出力するので、より具体的な客観評価を実現できる能力評価システムを提供できる。

【0020】また、請求項7に記載の発明では、情報の受付手段と、情報の記憶手段と、これら各手段の動作を制御する処理手段と、を備え、当該記憶手段が、文章を構成する自然文を単語に分解する基準となる基準単位単語を規定した分析用辞書記憶手段と、当該自然文を分解して得られた複数の単語を記憶する分解単語記憶手段と、当該文章に対応する分析値の基準分析値を記憶する基準分析値記憶手段と、当該分析値に基づき得られる乖離値についての比較基準を記憶する基準指標値記憶手段と、を有する能力評価システムにおいて、実行される能力評価プログラムを採用する。

【0021】そして、当該プログラムは、前記処理手段に、前記受付手段を介して自然文からなる文章を受け付けるステップと、前記分析用辞書記憶手段を参照して、前記基準単位単語に相当する複数の単語に前記自然文を分解するとともに、自然文を分解して得られた複数の単語を前記分解単語記憶手段に記憶するステップと、当該分解単語記憶手段が記憶した複数の単語を参照し、少な

くとも当該単語にかかわる数をカウント値として取得するステップと、当該カウント値に基づき複数の前記分析値を生成するステップと、生成した分析値と前記基準分析値記憶手段が記憶する基準分析値とに基づき、当該基準分析値に対する乖離の程度を示す複数の乖離値を生成するステップと、前記基準指標値記憶手段から読み出した基準指標値に基づき、生成した複数の前記乖離値同士の相対関係を比較するステップと、当該相対関係についての比較結果を出力するステップと、を実行させる。

【0022】本発明によると、分解された複数の単語に基づき得られる、所定の基準分析値に対する乖離値を用いて乖離値同士の相対関係を比較するので、客観的に文章内容の評価を行うことができる。ひいては、複数の文章について、それぞれ比較結果を出力することにより、文章同士の相対評価も容易に行うことができる。

【0023】また、請求項8に記載の発明では、請求項7に記載のプログラムにおいて、前記処理手段に、前記乖離値同士の相対関係を比較するステップとして、複数の前記乖離値に基づき、乖離値相互の相対関係を表すパターン指標値を生成するステップと、当該生成したパターン指標値と、前記基準指標値記憶手段から読み出した基準指標値とを比較するステップと、を実行させる。

【0024】本発明によると、乖離値の相対関係を表すパターン指標値を生成するので、より文章の持つあいまいさを加味した上で、文章のバランスについての確かな比較を行うことができる。

【0025】また、請求項9に記載の発明では、請求項8に記載の能力評価プログラムにおいて、前記処理手段に、前記相対関係として、一の乖離値を基準として他の乖離値との差を算出し、当該差の値に基づき各乖離値に対応するパターン指標値を生成するステップを実行させる。

【0026】本発明によると、一の乖離値を基準として相対関係をみるので、乖離値の相対関係がより明確となる普遍化ができ、また、比較パターン数を少なくすることにより、マッチング処理の高速化が図れる。

【0027】また、請求項10に記載の発明では、請求項7乃至請求項9に記載の能力評価プログラムにおいて、前記分析用辞書記憶手段は、少なくとも単語の同義語もしくは類義語のいずれかを規定する比較用辞書記憶領域を備えるものであって、前記処理手段に、前記自然文を分解して得られた複数の単語の中から、キーワードを抽出するステップと、当該キーワードを文章毎に比較し、前記比較用辞書記憶領域を参照して同義語または類義語の関係にあるか、もしくはキーワードが一致するか、の判断を行うステップと、当該判断の結果を出力するステップと、を実行させる。

【0028】本発明によると、分解された複数の単語から抽出されたキーワードの同義語や類義語まで比較可能なので、単語の意味内容まで加味した精度の高い文章内

容の評価を行うことができる。

【0029】また、請求項11に記載の発明では、請求項10に記載の能力評価プログラムにおいて、さらに、同義語または類義語の関係にあるか、あるいは、一致すると前記キーワード比較手段が判断したキーワードについて、前記分析値に基づき文章毎の使用頻度を比較するステップと、当該使用頻度の比較結果を出力するステップと、を実行させる。

【0030】本発明によると、キーワードの文章中で使用された頻度に基づく比較を行うので、文章中における書き手の単語の位置づけまでを含めた評価を実現する。

【0031】また、請求項12に記載の発明では、請求項7乃至請求項11に記載の能力評価プログラムにおいて、前記処理手段に、前記複数の乖離値に基づき複数の頭在能力評価値を生成するステップと、当該頭在能力評価値を出力するステップと、を実行させる。

【0032】本発明によると、乖離値に基づき頭在能力評価値を生成、出力するので、より具体的な客観評価を実現できる能力評価プログラムを提供できる。

【0033】

【発明の実施の形態】本発明にかかる能力評価を行う際の基礎となる、文章についての言語分析は、所定のテーマについて書かれた文章を評価することを主眼におくものであり、発明者による長年の研究・分析の結果から得られた経験的、統計的事実に基づく言語分析手法を能力評価手法に取り入れ、新しい能力評価システムおよび能力評価プログラムを具現化するものである。

【0034】発明者の研究によると、論文や社説など、各ジャンルにおいて一つのテーマを論じる文章として優秀と判断される文章を、50文章、100文章、500文章と分析を進め、文章に含まれる単語の出現回数に着目したところ、出現回数が多い単語は書き手の論旨の中心を表すキーワードとなっていることと、多くの単語が一回しか使用されないことが判明した。

【0035】例えば、文章中に100の単語が用いられている場合、出現回数の多い、論旨の本質を示す中心となるキーワードはおおよそ5つに過ぎず、おおよそ80の単語は一度だけ用いられているだけである。各ジャンル毎に比較すれば数値上で多少の差異は生じるが、このキーワードの文章に含まれる比率自体は、一のテーマを論じる文章である限り、同一のジャンルであれば60.0文字の文章でも、2000文字の文章でも基本的には変わりがなくとも分析結果として得られている。

【0036】さらに、単語の出現回数について言及すると、極論として、すべての単語が一度だけしか用いられないと、文章としては、とりとめのない言葉の羅列に過ぎず、論旨が浮かび上がってこないものとなる。一方、すべてが同じ単語となってしまえば、文章として成立しないということになる。すなわち、一つのテーマを効果的に相手に伝えるためには、単語の使用則として、繰

り返し用いられる単語が一定量あり、加えて、一度しか使われない単語がその周辺をとりまくように存在する必要があるというイメージが成立する。なお、出てきた単語数（重複を含む）を、出てきた単語の種類数で割った値（単語重複率）は、論文の場合は1.42、小説では2.01となるという分析結果も得られている。

【0037】上述した観点から、ある一定以上の割合で用いられるキーワードの割合を分析すると、例えば上述した100の単語が用いられている場合で説明すると、キーワードの数が5を超えてさらに多い場合には、主張しようとする事柄が多くなり、それぞれの論点の説明に十分な論証がおこなわれていない傾向が強い。逆に、キーワードの数が5を満たさず、さらに少ない場合には、主張すべき事柄への内容の集約が不足し、論旨を展開するにあたっての組み立てが不十分な傾向が強い。

【0038】他方、一度しか用いられていない単語群が80%を超えると、主張に対する論証として展開する方向が散漫となり、文章としての明確な方向性が表現しきれない傾向が強くなる。

【0039】以上説明したように、単語の重複という観点から見ると、文章の優劣は絶対量ではなくバランスであり、理想的な比率という基準で判断できるという知見を発明者は得ている。ここで、優秀な文章とは、文意が明確で、その論証が過不足なく行われており、押しつけや曖昧さがない文章を想定している。

【0040】このバランスという観点からの分析についてみると、品詞などの単語の属性の観点から見ても同様の結果を得る。例えば、名詞、形容詞、動詞、助動詞などの一般的な品詞区分で考えてみると容易に理解できる。名詞は、主語や目的語として用いられるが、自然文として、主語がなければ、動作主体がわからず意味不明となるし、目的語がなければ、動作対象などがわからず、同じく意味不明となる。また、形容詞などが少なければ、第三者に問題の状況やそれを取りまく環境などの詳細なイメージを十分に伝えられないし、逆に多すぎれば、与えられる情報が氾濫して読み手に伝えたいイメージがぼやけてしまう可能性が高い。そして、極論すると、動詞がないということは、自然文としての述語が存在しないことになり、少なくとも、主語に応じた比率で存在しなければ文章が成立しない。

【0041】単語の属性という観点からの分析を行ったことにより、属性の観点からもバランス良く用いられているか否かを、文章としての優劣を評価する基準とできるという知見を発明者は得ることができた。なお、本発明では、後述するように一般的な国文法の品詞種類とは異なる分析区分を採用しているが、この分析区分は、わかりやすい文章表現、文章構成という見地から重要となる、使用単語のバランスから文章を評価するという観点に基づき、どのような単語の属性を抽出すると、適切な数値評価ができる分析値が得られるかという分析実験を

繰り返して、経験的かつ統計的に得られた区分である。

【0042】さらに、上述したキーワードについて、より詳細に研究すると、文章中に二度以上用いられている単語は、上述したように全単語群の20%前後となるが、これらは文章の中心となっている単語群と、これを補足する単語群とからなる。統計的には、全単語群の5%前後が中心を表すキーワード群となり、全単語群の15%前後がこのキーワード群を説明する単語群となる。

【0043】また、1つの文章中で、書き手が無意識のうちに、重要な単語群とそうでない単語群を区別する傾向にあり、加えて、よく使われる単語の出現回数や、文意と直接関係のない単語群の存在する割合などに、書き手の能力が色濃く反映されることが判明した。したがって、1つの文章に用いられた単語群には、封じ込められた書き手の想いが表れるし、また、単語群の使い方から、読み手に対して自らが意図したとおりに伝えられるかという面での表現能力、説得能力、論理構成能力なども窺い知れることになる。

【0044】そして、この結果に基づくさらなる発明者の研究・分析により、単純なキーワードの数と他の単語の数とのバランスだけでなく、キーワードの使用頻度の観点から文章単位で分析することにより、文章の性質の段階的な評価、例えば、趣旨がよくまとまっている文章である、何かに固執している文章である、意味が分からない文章である、というそれぞれの段階を、分析結果から得られる数値の違いで評価することが可能であることがわかった。

【0045】特に、文章を普通名詞、固有名詞、動詞、形容詞という品詞区分の単語に着目して分解し、出現回数が最も高い単語を抽出し、文章毎にこの出現回数をすべての単語の出現回数の和で除した比率（以下、出現比率という）を比較してみると、文章の趣旨がまとまっていると思われる文章では、出現比率が概ね2.2~2.7となり、何かに固執している文章では、出現比率が概ね3.5以上となり、意味のわからない文章では、出現比率が概ね1.5以下になるということがわかった。検証として社説について分析してみると、出現比率は3前後となった。また、優劣取り混ぜた文章について無差別に検証してみると、サンプル数を増やすにつれて平均が3に集約する傾向が表れてきた。

【0046】ゆえに、文章のわかりやすさをバランスの面からみると、他人に特定の意味や意義を伝達するために文章を用いる場合、その内容について述べた文章の出現回数が最も高い単語の出現比率が3前後となるのが最も伝わりやすいといえ、この現象は、ある意味、日本語文化を示すものといえと推論できる。

【0047】以上の分析・研究の結果に基づく言語分析を基礎とする能力評価を実現するための能力評価システムおよび能力評価プログラムについて、以下に実施形態を説明するが、これらの実施形態は本発明の一例を説明

するに過ぎず、本発明の権利範囲を限定するものではない。また、説明において特記しなくとも、各手段は、当然に当業者として代替することができる手段を含むものとする。

【0048】〔第1実施形態〕

【0049】本発明にかかる能力評価システム1は、文章を用いて分析・評価を行うものであるが、大きく分けて3つの分析方法を具現化する。その分析方法とは、基準能力を備える所定のグループサンプルから求めた基準分析値に対する比較を行う基準分析方法、優秀だと誰もが認める文章をサンプルとして求めた基準分析値に対する比較を行う絶対分析方法、および、絶対分析で取得した基準分析値を用いて分析対象となる文章を相互比較する比較分析方法である。いずれの方法を実現するにせよ、所定の評価基準値に基づき、文章の分析・評価を行う手順を採用することになる。このような分析・評価を行う本発明の一実施形態を図1乃至図6に基づき説明する。

【0050】本実施形態では、特に、人材採用にあたって応募者の持っている能力を客観評価する一手段として、企業が設定した課題について応募者に論述させることにより得た文章を分析することにより、戦略遂行能力の有無を判断する材料を取得するために、能力評価システムおよび能力評価プログラムを用いる。なお、この場合に採用される分析方法は、基準分析方法や比較分析方法が想定される。

【0051】〔能力評価システムの構成〕図1は、本発明に係る能力評価システム1のシステム構成を示す機能ブロック図である。図1に示したように、能力評価システム1は、自然文からなる文章を受け付ける文章受付手段11と、当該自然文を単語に分解する基準となる基準単位単語を規定した分析用辞書記憶手段13と、分析用辞書記憶手段13を参照して、前記基準単位単語に相当する複数の単語に前記自然文を分解する自然文分解手段14と、自然文分解手段14が自然文を分解して得られた複数の単語を記憶する分解単語記憶手段15と、分解単語記憶手段15が記憶した複数の単語を参照し、少なくとも当該単語にかかわる数を前記文章についてのカウント値として取得するカウント値取得手段16と、カウント値取得手段16が取得したカウント値に基づき複数の分析値を生成する分析値生成手段18と、当該分析値の基準分析値を記憶する基準分析値記憶手段19と、基準分析値記憶手段19から読み出した基準分析値と、生成された前記分析値と、に基づき当該基準分析値に対する当該分析値の乖離の程度を示す複数の乖離値を生成する乖離値生成手段20と、当該複数の乖離値についての比較基準を記憶する基準指標値記憶手段21と、基準指標値記憶手段21から読み出した基準指標値に基づき、乖離値生成手段20が生成した複数の前記乖離値同士の相対関係を比較する乖離値比較手段22と、乖離値比較

手段22が行った比較結果を出力する出力手段23と、を主たる構成として備える。

【0052】そして、さらに、図1に示したとおり、能力評価システム1は、自然文分解手段14が自然文を分解して得られた複数の単語の中から、キーワードを抽出するキーワード抽出手段24と、キーワード抽出手段24が抽出したキーワードを文章毎に比較し、分析用辞書記憶手段13が備える、少なくとも単語の同義語もしくは類義語のいずれかを規定する前記比較用辞書記憶領域を参照して同義語または類義語の関係にあるか、もしくはキーワードが一致するか、の判断を行うキーワード比較手段26と、同義語または類義語の関係にあるか、あるいは、一致するとキーワード比較手段26が判断したキーワードについて、前記分析値に基づき文章毎の使用頻度を比較するキーワード使用頻度比較手段27と、乖離値生成手段20が生成した複数の乖離値に基づき複数の顕在能力評価値を生成する顕在能力評価値生成手段28と、を備える。

【0053】また、図2は、本発明にかかる能力評価システム1を実現するための概略的なシステム構成図である。図2に示したように、システム全体としては、情報の入力手段31、情報の表示手段32、情報の記憶手段33、外部装置との通信手段34、印刷手段35、およびこれら各手段の動作を制御する処理手段36とを有する構成を採用する。

【0054】図2において、入力手段31は、キーボードやスキャナなどの情報を入力するための手段である。表示手段32は、液晶ディスプレイやCRTなどの情報を表示するための手段である。通信手段34は、モデム、TA、ルータなど、外部装置との間で情報を受信し、送信するための手段である。印刷手段35は、電子情報を紙媒体などに出力して提供する手段であり、例えば、プリンタなどが該当する。なお、説明の容易のため、図2においてこれら入力手段31、表示手段32、通信手段34の数をそれぞれ一としたが、これに限らるものではない。

【0055】そして、記憶手段33は、プログラムやデータの格納領域を提供するとともに、プログラムの展開領域を提供する手段である。例えば、RAM、ROM、HDDなどの主記憶装置、補助記憶装置が挙げられる。ここで、図2に示したように、記憶手段33は、所定の記憶領域に文章記憶手段12と、分析用辞書記憶手段13と、分解単語記憶手段15と、パラメータ記憶手段17と、基準分析値記憶手段19と、基準指標値記憶手段21と、キーワード関連情報記憶手段25を含むものであり、各記憶手段の概要は以下のとおりである。

【0056】文章記憶手段12は、分析対象となる自然文を格納するための記憶領域を提供する。分析用辞書記憶手段13は、後述する分析用の辞書を格納するための記憶領域を提供する。分解単語記憶手段15は、後述す

る分析結果を格納するための記憶領域を提供する。パラメータ記憶手段17は、図1に示したカウント値記憶領域17a、分析値記憶領域17b、乖離値記憶領域17c、パターン指標値記憶領域17d、パターン比較結果記憶領域17e、顕在能力評価値記憶領域17fなど、分析に用いるための基礎データとなる数値や、演算結果として得られる数値などのパラメータを格納するための記憶領域を提供する。ここで、パラメータ記憶手段17は、これら各記憶領域を備えることにより、カウント値記憶手段、分析値記憶手段、乖離値記憶手段、パターン指標値記憶手段、パターン比較結果記憶手段、顕在能力評価値記憶手段としても機能するものである。基準分析値記憶手段19は、各文章の分析値を算出する際の基準となる基準分析値を格納するための記憶領域を提供する。基準指標値記憶手段21は、各文章から得られた乖離値を比較する際の比較基準を記憶する記憶領域を提供する。キーワード関連情報記憶手段25は、図1に示した、基準キーワード記憶領域25a、キーワード抽出結果記憶領域25b、判断結果記憶領域25cなどの、キーワードを抽出、比較判断した結果を格納するための記憶領域を提供する。

【0057】なお、説明の容易のため、図2において記憶手段33の数を一つとしたが、これに限られるものではなく、物理的に複数の記憶媒体から記憶手段が構成されるようにして、分散的に情報を記憶できるシステムとしてもよいことはいふまでもない。

【0058】処理手段36は、CPUなどの演算処理手段であり、上述した各種機器を制御する手段である。なお、説明の容易のため、図2において処理手段36の数を一つとしたが、これに限られるものではなく、物理的に複数の演算処理手段を用いて並列的に制御を行う構成としてもよい。また、処理手段36が仮想計算機などを介して、中間コード形式で提供されたプログラムを解釈、実行して処理を実現する構成としてもよい。

【0059】以下、図1に示した能力評価システム1を構成する各手段につき詳述する。

【0060】まず、図1における、文章受付手段11、自然文分解手段14、カウント値取得手段16、分析値生成手段18、基準分析値記憶手段19、乖離値生成手段20、乖離値比較手段22、キーワード抽出手段24、キーワード比較手段26、キーワード使用頻度比較手段27、顕在能力評価値生成手段28、および出力手段23は、能力評価システム1の処理手段36が所定のプログラムを実行することにより、実現され、その機能を発揮するプロセスである。なお、それらのプログラムは、通常、記憶手段33の所定の記憶領域に格納され、必要に応じて処理手段36が読み出し、メモリなどの記憶手段33のプログラム展開領域で実行され、各種機能を呈する。

【0061】ここで、文章受付手段11は、記載内容を

評価する分析対象である文章を受け付ける手段である。ただし、文章受付手段11が受け付ける分析対象の単位は、処理の効率を考慮して分析対象となる複数の自然文からなる文章を想定するものであるが、これに限られるものではない。例えば、設定した課題についての回答をアンケート形式などで収集した場合などは、回答として書かれた文章が一自然文単位となる場合があり得る。

【0062】また、文章受付手段11は、入力手段31を介して直接入力される情報を受け付けるという受け渡し形態に限られるものではない。例えば、能力評価システム1の外部から、通信手段34を介してテキスト形式などで自然文の電子情報を受け付けるようにしてもよいし、書き換え可能なMOやCD-RWなどの記録媒体などを介して受け付けるようにしてもよい。加えて、記憶手段33の他の記憶領域に格納されている文書データなどから分析対象を受け付ける形態を用いてもよい。

【0063】そして、文章受付手段11は、処理手段36の制御下で、受け付けた文章を文章記憶手段12に格納する機能も備える。これにより、能力評価システム1は、分析対象となる文章を受け付けた後、即座に分析するシーケンスを実行しなければならないという制約を受けることはなく、所定の数の文章を受け付けてから、任意のタイミングで分析するシーケンスを実行できる。

【0064】ここで、文章記憶手段12は、前述したように、記憶手段33の所定の領域に自然文を格納する手段であるが、その格納形態は、文章という単位で複数の自然文を格納するものに限られない。例えば、分析単位である文章との関連付けを維持しながら、一の自然文を独立したかたちで格納してもよいし、文章単位、自然文単位が混在したかたちで格納してもよい。ただし、複数の文章を分析処理し、比較するために、文章に固有の識別情報を付して、分析対象となる自然文をその識別情報に関連付けて記憶する手法を採用する。ここで付した識別情報は、後述する分析値、乖離値、パターン指標値、抽出されるキーワード、キーワードに関連する判断結果、顕在能力評価値などの情報とも関連付け、文章毎の能力評価情報の管理を実現する。

【0065】分析用辞書記憶手段13は、文章受付手段11が受け付けた自然文を単語に分解する際に参照する、基準となる基準単位単語を規定した分析用辞書として機能する手段である。具体的には、分析用辞書記憶手段13は分析単位となる基準単位単語を記憶する記憶手段であり、加えて、基準単位単語の属性に応じた分析区分を当該基準単位単語に付与し、相互に関連付けて記憶する記憶手段である。

【0066】ここで、基準単位単語の属性とは、一般的に単語の性質や特徴などをいい、例えば品詞の種類や意味に応じた分類などから判断される。そして、分析区分は、後述する評価値を、複数の観点から分析したものと意味づけるために複数生成される分析値を導出できるよ

うに、属性を考慮して定められる。

【0067】以下、分析用辞書記憶手段13に格納される辞書について、図3に基づき説明する。図3は、分析用辞書記憶手段13のデータ構造の一例を示す模式的構造図である。

【0068】図3に示したように、分析用辞書記憶手段13は、複数の独立した辞書から構成され、詳細には、体言辞書、用言辞書、接続詞辞書、助詞辞書、助動詞辞書、とともに、単語の語意の分析を念頭に、類義語・同義語辞書を備える。ここで、本実施形態では、類義語・同義語辞書の双方を備えたが、少なくとも単語の同義語もしくは類義語のいずれかを規定する比較用辞書を備えることにより、単語の完全一致を判断するのみの分析よりも高レベルの分析を可能にできる。

【0069】そして、分析用辞書記憶手段13は、上述した辞書以外に、分解に適した自然文、文節、単語レベルに修正することを主目的として、活用辞書、削除記号辞書、表記辞書、句点識別辞書などを備える。各辞書は、データベースとして構成され、所定の記憶領域に格納されるが、上述した基準単位単語を格納する体言辞書、用言辞書などは、分析区分に関連付けて基準単位単語を格納する際に、単語に付与された分析区分が利用され、関連付けられることになる。

【0070】ここで、分析用辞書記憶手段13の各辞書記憶領域に格納される基準単位単語に関連付けられる分析区分は、分解後の分析内容に応じて定められるものであり、本発明では、一般的な国文法をそのまま用いず、独自のアレンジを施した分析区分を採用する。

【0071】分析区分の具体例を表1および表2に示す。表1および表2に示したように、本実施形態では、分析用辞書記憶手段13に格納される基準単位単語の分析区分に階層構造を採用する。詳細には、表1は、第1階層が体言に区分される部分の詳細な階層構造を示した表であり、表2は、第1階層が用言と付属語に区分される部分の詳細な階層構造を示した表である。

【0072】表1および表2に示したように、基準単位単語は、第1階層として、体言、用言、付属語の3つの区分に分けらる。それぞれ第1階層内での区分を表す第1階層区分コードとして01、02、03というコードが付与される。

【0073】また、第2階層は、第1階層の体言、用言、付属語の下位区分という位置づけになる。具体的には、第1階層の体言は、普通名詞、固有名詞、代名詞、

数詞、接続詞、連体詞、感動詞に分けられる。第1階層の用言は、動詞、形容詞に分けらる。第1階層の付属語は、助動詞、助詞、連語に分けられる。第2階層のそれぞれの区分には、第2階層内での区分を表す第2階層区分コードとして、01から13までのコードが付与される。

【0074】第3階層では、表1および表2に示したように、第2階層内での一区分が、属性に応じてさらに下位区分に細分化される。代表的な例では、普通名詞は、普通名詞、成語、季語、時期、その他の区分に分けられ、第3階層区分コードとして、01、02、03、04、09というコードが付与される。また、固有名詞は、人名、社名、地名、国名、元号、歴史用語、その他の区分に分けられ、第3階層区分コードとして、01、02、03、04、05、06、09というコードが付与される。この第3階層の区分については、表1および表2に記載した構造となっており、上述した以外の区分および区分コードについては説明を省略する。なお、本実施形態における区分コードにより分析区分を特定する方式として、第3階層の分析区分を示す場合には、少なくとも第2階層区分コードとセットにして第3階層区分コードを用いる。

【0075】そして、表1および表2に示したように、本実施形態では、上述した分析区分に加え、基準単位単語の属性に鑑みて定められた抽出区分を規定する。ここで、抽出区分は、一括して抽出すべき分析区分の単語を特定するために設けられた便宜的な区分であり、第2階層区分の所定の区分に応じて、a、b、cの3つの区分が指定されている。例えば、キーワードとして分解した単語を選別する際に、区分aと区分bに該当する単語を抽出する際のインデックスとして用いる。

【0076】なお、これらの分析区分の階層構造データは、分解後に、分析区分に基づくカウントが行われる際の利便性を考慮して、記憶手段の所定の領域に、分析用辞書記憶手段13に格納される基準単位単語に関連付けられる分析区分を規定するデータとして格納される。格納の形態は、例えば、データベース構造などを採用し、各辞書の分析区分の階層構造に応じて、階層区分と分析区分コードとの対応付けを規定するとともに、高速検索のための検索インデックスとして抽出区分との対応付けも規定する。

【表1】

第1階層	(コード)	第2階層	(コード)	第3階層	(コード)	抽出区分
体言	01	普通名詞	01	普通名詞	01	b
				成語	02	b
				季語	03	b
				時期	04	b
				その他	09	b
		固有名詞	02	人名	01	b
				社名	02	b
				地名	03	b
				国名	04	b
				元号	05	b
				歴史用語	06	b
				その他	09	b
		代名詞	03	人称代名詞	01	a
				指示代名詞	02	a
				その他	09	a
		数詞	04	年月日	01	a
				金額	02	a
				単位	03	a
				その他	09	a
		副詞	05	状態副詞	01	a
				程度副詞	02	a
				陳述副詞	03	a
				その他	09	a
		接続詞	06	順接	01	a
				逆接	02	a
				因果	03	a
				並列	04	a
				添加	05	a
				補足	06	a
				選択	07	a
				転換	08	a
				その他	09	a
		連体詞	07			a
		感動詞	08	感動	01	a
				呼びかけ	02	a
				応答	03	a
				掛け声	04	a
				挨拶	05	a

【表2】

第1階層 (コード)	第2階層 (コード)	第3階層 (コード)	抽出区分
用言	02	自動詞 01	b
		他動詞 02	b
		一般動詞 03	b
付風語	03	形容詞 10	b
		使役 01	c
		受身・尊敬・自発・可能 02	c
		打消 03	c
		推量・意志・勧誘 04	c
		打消推量・意志 05	c
		過去・完了・存続 06	c
		希望 07	c
		丁寧 08	c
		様態 09	c
		伝聞 10	c
		比例・例示・不確定断定 11	c
		断定 12	c
		推量 13	c
		丁寧断定 14	c
		格助詞 01	c
		接続助詞 02	c
		副助詞 03	c
		係助詞 04	c
		終助詞 05	c
		間投助詞 06	c
		連語 13	c

【0077】ここで、区分aは、内容を修飾または補足する単語群をまとめた区分を意味する。区分bは、内容を具体的に示す単語群をまとめた区分を意味する。区分cは、その単語だけでは意味をなさない単語群をまとめた区分を意味する。

【0078】以上説明した分析区分と関連付けられた基準単位単語を格納する各辞書およびその他の辞書について詳述する。

【0079】体言辞書記憶領域は、一般的な国文法では普通名詞、固有名詞、代名詞、数詞（主に単位など）、副詞、連体詞、感動詞に分類される基準単位単語を格納する。ここで、本発明に特徴的なことは、一般的な国文法では体言に分類される接続詞について、独立したかたちで接続詞辞書を設けていることと、連語を体言辞書に登録していることである。

【0080】これは、接続詞辞書が、実際に文章表現内容の分析を行う際に高い頻度で利用されることに鑑み、体言辞書に含めず独立させて設ける構成を採用したものであり、自然文を分解して単語にする際に用いることは離れて、異なる言語分析を行う構成を付加的に備えた場合などを想定し、文章表現内容の分析に兼用しやすいようにしたものである。また、連語も二つ以上の単語が結合してはいるが、ひとまとまりの観念を表す語であるので、体言辞書に格納している。

【0081】用言辞書記憶領域は、一般的な国文法では動詞、形容詞に分類される基準単位単語を格納し、同様に、助詞辞書記憶領域は、助詞に分類される基準単位

単語を、助動詞記憶領域は、助動詞に分類される基準単位単語をそれぞれ格納する。

【0082】また、活用辞書記憶領域は、一般的な国文法では動詞、形容詞に分類される基準単位単語の活用として語尾変化のパターンなどを格納しており、自然文の分解に際して、その単語が用言であるか否かを認定する際に用言辞書と併用される。

【0083】句点識別辞書記憶領域は、自然文の末尾を認識するために、句点（。）を格納する。ただし、必要に応じて、句点識別辞書記憶領域は、さらに、読点（、）や疑問符（？）、感嘆符（！）などを格納してもよい。

【0084】表記辞書記憶領域は、電子文書などで特殊記号に分類される自然文中の記号を変換するために、変換元情報と変換先情報を関連付けて格納する。例えば、百分率記号（％）と「パーセント」という単語、ドル記号（\$）と「ドル」という単語、円（¥）と「円」または「エン」などを関連付けて格納する。

【0085】削除記号辞書記憶領域は、分解や分析に関係しない記号で、削除しても問題がない記号を格納する。例えば、文章中で見出しなどに使う「○」「●」「◎」「§」「※」などの記号が挙げられる。

【0086】類義語・同義語辞書記憶領域は、後述するキーワードの語意を考慮した比較分析を行うために用いられる比較用辞書記憶領域であり、同義語や類義語の関係にある単語を相互に関連付けて格納する。

【0087】感想単語辞書記憶領域は、自然文を単語に

分解する段階では使用しないが、後述する中心キーワード比率や中心補完単語比率などの分析値を生成する際に必要となる感想単語を格納する。詳細には、感想単語辞書記憶領域は、本発明独自の区分として、感想単語という分析用の特別区分を設け、これに該当する単語の数を差し引いた分析値を生成する処理を行うために、その感想単語に該当する基準単位単語を格納し、一致判断の際の参照辞書とする辞書記憶領域である。

【0088】ここで、感想単語に該当するのは、「私は…と思う」と表現される場合の、「私」「思う」など、自然文制作者の主観的な感想として位置づけられる、主体を自分として述べるときに用いられる単語である。例えば、感想名詞に該当する単語として、「自ら」「自分」「私」「僕」「当社」「我社」など、感想動詞に該当する単語として、「思う」「できる」「対する」「考える」「感じる」などが挙げられる。

【0089】なお、本実施形態では、体言辞書記憶領域、用言辞書記憶領域、接続詞辞書記憶領域、など、概ね分析区分と一致する構成として、自然文を単語に分解する際の参照の高速化を図れるようにしたが、これに限られるものではない。例えば、分析区分と対応付けられ、自然文の分解に際して基準単位単語として参照する分析手順を実現できるものであれば、独立した単語群を形成するように格納されたデータベース構造に代えて、単語がその分析区分や品詞種別に基づきまとめられることなく混在したデータベース構造を採用してもよい。

【0090】また、必ずしも単一の分析用辞書記憶手段13が、図3に示したように、複数の辞書を統括して記憶する構成に限られるものではなく、複数の記憶手段から分析用辞書記憶手段13が構成され、複数にグループ化された辞書群を分散して記憶する構成を採用してもよい。また、辞書も上述した機能別の辞書を複数備える構成に限るものではなく、例えば、用言辞書と活用辞書を統合した複合機能を果たす辞書を構成要素としてもよい。

【0091】続いて、自然文分解手段14は、分析用辞書記憶手段13に格納されている基準単位単語を参照して一致を判別し、基準単位単語に相当する複数の単語に自然文を分解する手段である。具体的には、自然文分解手段14は、句点識別辞書記憶領域を参照しながら、文章記憶手段12に格納された分析対象となる自然文をセンテンス単位で読み出し、読み出した自然文について、所定の文字数単位で自然文の先頭から順番に文字列を抜き出し、分析用辞書記憶手段13の体言辞書記憶領域、用言辞書記憶領域、接続詞辞書記憶領域など、所定の順序で各辞書を参照し、辞書内に格納された基準単位単語と比較して分解される単位となる単語を特定することにより単語に分解する手段として機能する。

【0092】また、自然文分解手段14は、複数の単語に自然文を分解するに際し、分解した単語に分析区分を

関連付ける手段としても機能する。具体的には、分析用辞書記憶手段13に格納される基準単位単語に関連付けられた分析区分を、分解した単語に関連付けるべき分析区分として用いることにより、分解した単語の属性に応じた分析区分を付与する手段として機能する。

【0093】加えて、自然文分解手段14は、自然文を分解した結果である単語群を分解単語記憶手段15に格納する。ここで、分解された単語の格納の形態は、分解すべき単語として抽出したことに誤りがないことが確定した単語を順次格納する形態としてもよいし、自然文単位で分解処理が完了する毎にまとめて分解結果を格納する形態を採用してもよい。

【0094】なお、自然文分解手段14が実現する、文章記憶手段12に格納された自然文を単語に分解するプロセスは、上述の態様に限られるものではなく、少なくとも単語とそれに応じた分析区分を関連付けて分解格納することができるものであれば足りる。

【0095】分解単語記憶手段15は、前述したように、自然文分解手段14が自然文を複数の単語に分解して得られた単語を記憶する手段であるが、自然文を分解して得られた単語と、その単語の属性に応じた分析区分とを関連付けて記憶する。ここで、分解された単語は、前述したような分析対象となる文章に固有の識別情報などを用いて、文章毎に独立したかたちで記憶される。

【0096】分解単語記憶手段15の単語の記憶形態は、自然文の先頭側から順に分解した単語を格納し、その単語に該当する分析区分のフラグを立てる形態や、単語と分析区分に対応した区分コードとを対応付けて格納する形態、分析区分毎に分けて単語を格納する形態などを採り得る。また、分解した単語の中に、同一の単語が複数含まれている場合は、重複させて格納してもよいし、その単語に重複数を関連付けて格納することにより重複させないようにしてもよい。

【0097】なお、分解単語記憶手段15は、自然文分解手段14が自然文を分解して取得した単語群などの、直接的な一次的結果を記憶する機能に限定されるものではない。分析対象毎に、対応する追加的な結果、例えば、一次的結果として得られた単語に、後述する同頻度数データやキーワードと認定されたことを示す識別フラグなどの二次的結果を関連付けて記憶する機能を付加してもよい。

【0098】カウント値取得手段16は、分解単語記憶手段15が記憶した複数の単語を参照し、少なくとも分解した前記単語にかかわる数を前記文章についてのカウント値として取得し、取得したカウント値をパラメータ記憶手段17のカウント値記憶領域17aに格納する手段である。ここで、「少なくとも」としたのは、自然文分解手段14が直接的に分解した単語数のみをカウントするのではなく、後述するように文字数や自然文の数（以下、センテンス数と称する）をカウントする手段と

しても機能し得るからである。

【0099】具体的には、カウント値取得手段16は、上述したような形態で分解単語記憶手段15が格納する分解された単語群に基づき、単語に関連付けられた分析区分を検索インデックスとするなどしながら単語の数をカウントし、そのカウント値を取得する。

【0100】カウント値記憶領域17aへのカウント値の記憶形態の一例として、図4にカウント値データベースのデータ構造図を示す。図4の各数値は、設定された課題に対して論述した文章を分析対象とした場合に取得されるカウント値を表す。

【0101】ここで、分解単語記憶手段15に格納された単語群に基づくカウント値の例としては、第1階層区分、第2階層区分、第3階層区分など、各区分が付与された単語の総数を区分毎にカウントしたカウント値が挙げられる。図4に示したカウント値として得られるデータ項目のうち、第1階層区分の体言（第1階層区分コード01）に該当する単語の数をカウントして得られた体言数、同様に、用言（第1階層区分コード02）に該当する単語の数をカウントして得られた用言数、付属語（第1階層区分コード03）に該当する単語の数をカウントして得られた付属語数などがこれに該当する。

【0102】ただし、カウント値取得手段16が取得するカウント値は、単に特定の分析区分に該当する単語の総数に限られるものではない。例えば、その分析区分に該当する単語でも、重複する同一単語はひとまとまり（一つ）としてカウントして求められる単語の種類数（以下、出現単語数と称する）が挙げられる。なお、ひとまとまりとしてカウントする手法としては、すでにカウント済みの単語と一致を判別し、一致する場合にはカウントをスキップなどして、空カウントとするなどの手法を採用する。

【0103】また、本実施形態では、カウント値取得手段16は、分析区分という枠内に該当する単語のみをカウント対象とするものではなく、その分析区分をまたがって単語数をカウントしてもよい。

【0104】例えば、分析対象となる文章中出现した回数（以下、この出現回数を頻度数と称する）が同じである単語の数をカウントして求められる頻度数が同じ単語の数（以下、同頻度数と称する）、さらには、頻度数が一である単語（分析対象文章中で一度しか出現していない単語）の数（以下、1頻度単語数と称する）、などをカウント値としてカウント値取得手段16が取得する態様も採り得る。

【0105】さらに、カウント値取得手段16は、単語の数のカウント値だけでなく、必要に応じて、分解単語記憶手段15のみならず、文章記憶手段12なども参照し、対象となる自然文の総文字数や、センテンス数、各分析区分に該当する単語の文字数を合計した数、などについてもカウントし、そのカウント値を取得する。ま

た、後述する分析値生成手段18により生成された分析値などに基づく再カウントなども、必要に応じて行う。

【0106】加えて、カウント値記憶領域17aには、分解した単語の数をカウントして得られたカウント数から直接得られるカウント値のみならず、場合によっては、演算処理などを施して間接的に得られるカウント値をも記憶する構成としてもよい。

【0107】そして、上述したように、分析区分に着目して単語の数をカウントしたカウント値を格納したのと同様に、それ以外のカウント値についても、パラメータ記憶手段17に設けられたカウント値記憶領域17aに格納する。

【0108】なお、図4に示したカウント値のうち、キーワード数や感想単語数は、所定のカウント値の処理がおこなわれた後、分析処理を挟んで、再度、必要に応じてカウントされるカウント値であり、詳細は後述する。

【0109】分析値生成手段18は、カウント値取得手段16が取得したカウント値に基づき複数の分析値を生成する手段であるとともに、パラメータ記憶手段17の分析値記憶領域17bに対して、算出した分析値を格納する手段としても機能する。具体的には、分析値生成手段18は、パラメータ記憶手段17のカウント値記憶領域17aに格納された各種カウント値を用いて分析値を生成する。また、分析値の生成にあたっては、分析値生成手段18は、記憶手段33の所定の領域に設けられた不図示の分析値算出則記憶領域に格納された分析値算出則を参照して算出する。

【0110】ここで、本実施形態における分析値とは、カウント値に基づき生成された値をいうが、カウント値を加工して得られた二次的な数値を用いて生成された値のみに限られるものではなく、カウント値記憶領域17aに格納されたカウント値をそのまま用いて生成された値も含む。

【0111】さらに、分析値は、上述した単語の数に基づくカウント値のみにより算出されなければならないものではなく、総文字数やセンテンス数などに基づくカウント値と、単語の数に基づくカウント値との双方を用いて生成されるものであってもよいことはいうまでもない。

【0112】本発明にかかる分析値の代表的なものを表3に示す。表3には、人の能力を文章の内容に基づき評価する際の判断項目となる判断基準能力を表す分析項目内容、分析項目内容に対応する分析値、および、その分析値の算出則を記している。

【0113】表3に示したように、文章の評価、書き手の能力評価の一要素として、例えば、文章整備力なら総頻度比率を評価指標の基礎データとして用いることができる。また、その算出則は総頻度数を計算文字数で除すというものである。なお、算出則に用いられる各要素の詳細は後述する。

【表3】

分析項目内容	対応分析値	分析値算出則
文章整備力	総頻度比率	総頻度数÷計算文字数
体言率	体言率	体言数÷計算文字数
用言率	用言率	用言数÷計算文字数
付属語率	付属語率	付属語数÷計算文字数
語彙力	除数詞名詞率	除数詞名詞数÷計算文字数
テーマ設定力	中心キーワード比率	中心キーワード数÷計算文字数
テーマ説明力	中心補完単語比率	中心補完単語数÷計算文字数
テーマ補足力	中心1頻度単語比率	中心1頻度単語数÷計算文字数
単語構成力	出現単語比率	出現単語数÷計算文字数
センテンス主張力	センテンス頻度平均値	総単語重量値÷センテンス数
センテンス構成力	センテンス比率	計算文字数÷センテンス数
主張力	第1単語頻度比率	第1単語数÷計算文字数
主張補足力	第2単語頻度比率	第2単語数÷計算文字数
主張構成力	第3単語重複率	(第1～第3単語数)÷3
論理展開力	同頻度数比率	同頻度数÷計算文字数
論理強制力	単語回帰値	(関数：最小自乗法使用)
行動表現力	センテンス用言率	用言数÷センテンス数

【0114】ここで、表3の分析項目内容に記載した各能力に対応付けられた分析値は、発明者の長年の研究および分析から得られた経験に基づき、文章の内容を評価する際の各能力を数値評価するに適していると認定されたものである。付言すれば、表3に示したように、単語の品詞属性等に基づく分析区分毎に、当該単語に関連するカウント数を用いて普遍化した比率が分析値の主流を占めるが、センテンスに関連するカウント数を用いて普遍化した比率も分析値となる。分析値として得られた数値を絶対比較しても、その能力に関する他の文章との優劣を判断することが可能である。なお、表3に示した分析値および分析値算出則は一例であり、それ以外の観点から算出される分析値を用いた評価を行う場合でも、好適に本発明を用いることができることはいうまでもない。

【0115】そして、本実施形態においては、表3に示したような分析値算出則記憶領域への分析値算出則の格納形態としては、プログラム関数もしくはモジュールの演算機能のかたちで記憶手段33に格納される構成を採用しており、分析値生成手段18が記憶手段33の所定の記憶領域に格納されたプログラム関数またはモジュールを利用して分析値を生成するという方式を採る。

【0116】ただし、分析値算出則の格納は、上述したように、関数として繰り返し用いることができるようにサブルーチンの格納する形態に限られるものではなく、シーケンス制御の手順として、重複して逐次規定され格納される形態であってもよい。

【0117】ゆえに、分析値生成手段18は、詳細には、カウント値記憶領域17aと記憶手段33に設けられた不図示の分析値算出則記憶領域とを参照し、カウント値記憶領域17aから読み出したカウント値と、分析値算出則記憶領域から読み出した分析値算出則と、に基づき、分析値を生成する手段として機能する。

【0118】続いて、分析値生成手段18により生成される分析値について詳述する。上述したように、分析値には、カウント値記憶領域17aに格納されたカウント値をそのまま用いて生成された分析値と、表3に示した分析値などのように、カウント値を加工して得られた二次的な数値を用いて生成された分析値がある。なお、表3においては分析値算出則の要素となっている計算文字数なども、カウント値に基づく分析値と位置づけてよい。

【0119】まず、直接的にカウント値を用いて得られる分析値としては、表3に示された総頻度数、除数詞名詞数やセンテンス用言率などが挙げられる。具体的に、総頻度数は、分析対象となる複数の自然文に含まれる分析区分の内、抽出区分aおよびbに該当する分析区分毎の単語の数をカウントして得られた総和をいう。また、除数詞名詞数は、第2階層区分が普通名詞、固有名詞、代名詞に該当する単語の数の総和から数詞に該当する単語の数を差し引いた単語数をいう。そして、センテンス用言率は、用言数をセンテンス数で除した数値をいう。

【0120】ここで、本実施形態において、総頻度数を

すべての単語の数にせず、抽出区分a、bに限ったのは、経験則から、抽出区分cに該当する単語の数を除外して分析を行うことにより、専門家が経験的によい文章であると判断した場合の人的分析評価に沿った、数値分析評価が得られると判明したからである。

【0121】一方、カウント値を加工して得られた二次的な数値を用いて生成される分析値としては、総頻度比率、体言率、などの比率または率で表される、計算文字数でカウント値等を除して求められる分析値や、中心キーワード比率、中心補完単語比率、中心1頻度単語比率、第1単語頻度比率、第2単語頻度比率、第3単語重複率のように、カウント値を用いて算出したパラメータを計算文字数で除した分析値などが挙げられる。

【0122】ここでいう計算文字数は、例えば固有名詞の文字数の多寡（外国の地名などの有無）や、対象となる文章自体の多寡など、物理的または表面的な要素により分析値が大きく影響を受けることがないように導入したパラメータであり、本発明に特徴的な概念である。

【0123】本実施形態における計算文字数の算出則は、（計算文字数）＝（分析対象となる文章の総文字数）×（総頻度数）／（総頻度数でカウントされた分析区分に該当する単語の文字数の総和）である。ゆえに、比率を算出する際の分母となるため文字数と称したが、単語の数である総頻度数と、文字数から得られた無次元係数とから得られる頻度数とを用いて、分析値を生成しているものであり、これにより、文字量の影響を極力排除せんとするものである。

【0124】より具体的には、冗長でわかりにくい文章と簡明で短い文章とを比較した際に、その文字量により左右されにくい分析値を得ることができるようにするため、単語の数に対して文字数に基づく補正をかけ、分析値の普遍化を実現する普遍化パラメータとして上述の計算文字数を選択したものである。

【0125】また、中心キーワード比率、中心補完単語比率、中心1頻度単語比率などは、分解した単語のうちキーワードと認定される単語数が重要なパラメータとなる。

【0126】先にも触れたが、キーワードとは、同頻度件数を同頻度のパターン数で除した同頻度比率が0.7以下の単語とされる。ここで、同頻度件数とは、特定の出現回数につき、その出現回数が同じである単語の種類数をいう。例えば、出現回数が4回の単語が3種類あったら、頻度数が4の同頻度件数は3とされる。また、同頻度のパターン数とは、同頻度と認定された単語群の数をいい、具体的には、出現回数が同じである単語群がいくつできるかをカウントしたものである。例えば、出現回数が1回の単語群、出現回数が2回の単語群、出現回数が3回の単語群、出現回数が4回の単語群、出現回数が6回の単語群、出現回数が12回の単語群がある場合には、同頻度のパターン数は6となる。このとき、頻度

数が4の単語群については、前述したように同頻度件数が3であるので、同頻度比率が0.5となり、キーワードとして認定されることになる。なお、参考までに、優秀な文章とされる社説のうちの1つについて分析した結果をあげると、頻度数12が同頻度件数1件、頻度数6が同頻度件数1件、頻度数4が同頻度件数3件、頻度数3が同頻度件数5件、頻度数2が同頻度件数18件、頻度数1が同頻度件数178件、となり、頻度数12、頻度数6、頻度数4までに該当する単語群がキーワードとして認定された。

【0127】このキーワード認定のために、分析値生成手段18は、分解された単語もしくは単語群がキーワードに該当するか否かを、カウント値記憶領域17aに格納された各カウント値に基づき算出した同頻度比率が0.7以上か否かにより判断する機能も備える。このとき、その判断結果に基づき、キーワードに該当する単語数（以下、キーワード数と称する）をカウント値取得手段16がカウントし、さらに、そのカウント値に基づき、分析値生成手段18が分析値を生成する。

【0128】具体的には、以上説明したキーワードに該当する単語数に基づき生成される分析値が、補完単語数であり、キーワード数から、前述した一頻度単語数を差し引いた数値をいう。

【0129】さらに、分析値生成手段18は、キーワードに該当すると判断したキーワード群や補完単語に該当すると判断した補完単語群の中に、感想単語が存在するか否かを判断する機能を備える。具体的には、前述した分析用辞書記憶手段13に格納されている感想単語辞書記憶領域を参照して、その感想単語として記憶されている基準単位単語との一致をみることにより判断する。

【0130】分析値生成手段18は、この感想単語の認定に基づき、中心キーワード数として、キーワード数から感想単語数を差し引いた数値を分析値として生成する。また、中心補完単語数として、補完単語数から感想単語数を差し引いた数値を分析値として生成する。さらに、中心一頻度単語数として、一頻度単語に該当する単語から感想単語を差し引いた数値を分析値として生成する。その上で、分析値生成手段18は、これら、中心キーワード数、中心補完単語数、中心一頻度単語数をそれぞれ計算文字数で除して、中心キーワード比率、中心補完単語比率、中心1頻度単語比率を生成し、分析値として分析値記憶領域17bに格納する。

【0131】ただし、本実施形態におけるキーワードの認定方法として、上述した同頻度件数を同頻度のパターン数で除したものを同頻度比率としたが、キーワードの概念は、前述したように、ある一定回数以上の出現頻度で使用されている単語というものであり、この区別を行うための指標の一例として上述の同頻度比率を用いた構成を採用したものである。したがって、キーワードを認定するための基準となる頻度を定められるものであれ

ば、これに限られるものではなく、キーワードを認定するための出現頻度のしきい値として、例えば、同頻度件数を総同頻度件数で除した値を同頻度比率として用いてもよい。

【0132】他の分析値としては、第1単語頻度比率、第2単語頻度比率、第3単語重複率が挙げられる。ここで、第1単語、第2単語、第3単語、第4単語は、カウント値取得手段16が取得したカウント値に基づき、頻度数（出現回数）の多い単語を上位から4つ抽出して特定する。ゆえに、それぞれの単語の頻度数は、第1単語数、第2単語数、第3単語数、第4単語数、として取り扱われ、各々を計算文字数で除した値が第1単語頻度比率、第2単語頻度比率、第3単語頻度比率、第4単語頻度比率という分析値になる。また、第3単語重複率とは、第1単語、第2単語、第3単語とされた単語について、その種類数の和をとり、3で除した値をいう。

【0133】そして、各分析値は、表3に示したように分析対象の文章の内容を表す数値となる。具体的には、文章整備力を表す数値として総頻度比率が、語彙力を表す数値として除数詞名詞率が、テーマ設定力を表す数値として中心キーワード比率が、テーマ説明力を表す数値として中心補完単語比率が、テーマ補足力を表す数値として中心1頻度単語比率が、単語構成力を表す数値として出現単語比率が、センテンス構成力を表す数値としてセンテンス比率が、センテンス主張力を表す数値としてセンテンス頻度平均値が、主張力を表す数値として第1単語頻度比率が、主張補足力を表す数値として第2単語頻度比率が、主張構成力を表す数値として第3単語重複率が、論理展開力を表す数値として同頻度数比率が、論理強制力を表す数値として単語回帰値が、行動表現力を表す数値としてセンテンス用言率が採用される。

【0134】ここで、センテンス頻度平均値とは、各センテンスの区分bに該当する単語について、その単語重量値の総和をとり、そのセンテンス毎に得られた単語重量値の総和をさらに合計した値を、センテンス数で除した値である。ここでいう単語重量値は、単語の頻度比率であり、出現回数（頻度数）を計算文字数で除したものである。

【0135】また、回帰値とはデータ分布を一つの直線に表したときの傾きを表す値であるが、単語回帰値の本来の意義は、すべての単語の頻度比率を分布図に表したときの直線の傾きをいう。ただし、本実施形態では、近似的に傾きを取得する構成を採用している。具体的には、縦軸のパラメータとして理想値を用い、横軸のパラメータとして分析対象文章の分析値をとり、理想値と合致する場合には45度の傾きをなす分布直線を想定する。そして、単語の頻度比率としては、第2単語頻度比率を第1単語頻度比率で除した値、第3単語頻度比率を第1単語頻度比率で除した値、第4単語頻度比率を第1単語頻度比率で除した値の総計3つの値を用いて、これ

らのパラメータを用いて傾きを求める。なお、第1単語頻度比率で除したのは、回帰値を得るための傾きを単調増加する直線に近似できるようにするためである。

【0136】詳細には、縦軸には、平均値±標準偏差の範囲に入ったときの値から算出された再平均値を用いて算出した理想値をとるとともに、横軸には、各分析対象文章の現実値をとり、これら理想値と現実値の3対のデータから最小自乗法により傾きを算出する。ここで、最小自乗法による算出には、公知のプログラム関数などを用いる。また、単語回帰値の傾きが表す傾向としては、傾きが45度よりもなだらかな場合には、冗長であいまいな表現となっており、45度より急な場合には命令張の断定的な表現となっているということが、経験的に得られている。

【0137】基準分析値記憶手段19は、文章に基づき得られる分析値の基準となる基準分析値を記憶する手段であり、文章についての言語分析・能力評価に先だって、評価基準となる数値評価基準分析値を設定する機能を果たす。

【0138】具体的には、比較分析を行う場合には、基準分析値記憶手段19は、優秀だと各ジャンルで認められる文章をサンプルとして求めた基準分析値を記憶する。この場合の基準分析値は、数百から数千のサンプルを分析し、その分析値の平均を基準分析値とする。これは、文章を分析するという本質から考えると、基準分析値が、日本語という文化を反映したものであることが望ましいといえるからであり、前述したバランスという観点から、日本の文化として集束する値を求めて基準分析値とするものである。したがって、比較分析の場合には、社説などの優秀な表現技法を実践する文章から基準分析値を求め、それを用い、これらの優秀な文章から得られた基準分析値を事前に基準分析値記憶手段19に記憶させておく構成とする。

【0139】一方、基準分析を行う場合には、例えば、人材を確保したい部署における優秀なスタッフに設定された課題について論述してもらった文章を用い、これを基準となるグループサンプルとして分析値を求め、平均値を基準分析値とする。基準分析の場合には、このグループサンプルとして得られた基準分析値を基準分析値記憶手段19が記憶するので、企業の社風や部署の特性などを色濃く反映した基準分析値が規定されることになる。

【0140】付言すれば、比較分析または基準分析を行う場合の基準分析値の取得に際しても、上述した文章受付手段11、分析用辞書記憶手段13、自然文分解手段14、分解単語記憶手段15、カウント値取得手段16、分析値生成手段18を用いることにより、好適に分析値が得られ、処理手段36により、その分析値の平均値を取得することができる。

【0141】ただし、ここで採用される平均値は、例え

ば、処理手段36が所定のプログラムに基づき実現する不図示の基準分析値生成手段により、複数の文章毎に生成された複数の分析値を用いて標準偏差を算出し、算出した標準偏差を用いて再度分析値をフィルタリングして抽出した分析値集団の平均値などが好適である。また、前述した分析値生成手段18が、複数種類の分析値を生成する場合には、不図示の基準分析値生成手段は、その種類毎に生成された分析値を用いて基準分析値を生成する。

【0142】なお、不図示の基準分析値生成手段が基準分析値を生成する際のフィルタリングとは、分析対象となる文章毎に生成された各分析値を用いて、複数の分析値について平均値と標準偏差を算出し、その上で、(平均値)±(標準偏差)の範囲に入っていない分析値を除外し、再度、平均値を求める(以下、再平均値と称する)。この再平均値を基準分析値、すなわち理想値として位置づけ、文章の分析を行う際の数値基準として用いる手順を採用することがより好ましい。

【0143】ただし、本発明においては、分析対象が文章という個人差が大きくなりやすいものを分析対象としているため、標準偏差を用いてばらつきの度合いが大きい標本を削除することにより、外乱的な要素を排除した評価基準を得る構成を採用してもよい。もっとも、多少の外乱的な要素を含む基準としても良い場合には、直接的に複数の分析値から得られる平均値を、上述したような数値操作を行うことなく、基準分析値としてもよい。

【0144】なお、上述したような基準分析値生成にあ

$$(\text{乖離値}) = \frac{B(\text{分析値} \times A - 1)}{C} \quad \dots (1)$$

【0148】なお、本実施形態においては、これらの補正係数を基準分析値と関連付けて基準分析値記憶手段19が格納する構成としたが、係数毎に必要なに応じて処理手段36が乖離値生成手段20の一機能として算出する構成としてもよいし、基準分析値と分けて各係数値を記憶手段33の所定の領域に格納する構成としてもよい。その構成に応じて、既に格納された係数を用いて乖離値生成手段20が乖離値を生成する場合には、基準分析値記憶手段19から読み出した基準分析値から算出した係数を用いることで、間接的に基準分析値に基づくかたちで乖離値を生成することになるし、一連の処理シーケンスとして乖離値生成手段20が係数を算出するとともに、算出した係数を用いて乖離値を生成する場合には、処理シーケンス全体から見ると直接的に基準分析値に基づき乖離値を生成することになる。

【0149】以下に、各係数の意義とともに、各係数の算出手順について参考までに説明する。まず、補正係数Aは、分析値毎に桁数異なるため、分析値を指標として用いる際に値を揃えて普遍化しやすいように、分析値を修正する係数である。式(1)において、原則として分析値に補正係数Aを乗じた値が、1付近を中央値とする

* たったの基準分析値算出則は、記憶手段33の所定の領域に設けられた不図示の基準分析値算出則記憶領域に、プログラム関数またはモジュールなどのかたちで格納されるものである。

【0145】乖離値生成手段20は、基準分析値記憶手段19から読み出した基準分析値と分析値生成手段18が生成した分析値とに基づき、読み出した基準分析値に対する分析値の乖離の程度を示す乖離値を生成する手段である。

【0146】ここで、乖離値を算出するための算出則を次式(1)に示す。この算出則は、プログラム関数やモジュールなどのかたちで、前述した分析値算出則や基準分析値算出則と同様に、記憶手段33の所定の領域に設けられた不図示の乖離値算出則記憶領域に格納される。

【0147】式(1)において、Aを補正係数、Bを段階係数、Cを分散係数と称する。図5に、前述した基準分析値とともに、各係数値を格納したデータベースのデータ構造の一例を示す。ここで、図5に示したデータベースは、比較分析を行う場合の基準分析値となる再平均値と、基準分析を行う場合の基準分析値とを両方格納し、さらに、職種毎の基準分析で用いる基準分析値となる再平均値を複数備える構成を採用する。また、このデータベースは、基準分析値記憶手段19であるとともに、上述した乖離値算出用の係数を記憶する所定の係数記憶領域として機能するものである。

【数1】

ように補正係数Aを定める。ゆえに、この補正係数Aは、基準分析値で1を除すという手順で得られるものであり、基準分析値に基づく乖離値の生成を行う処理の根幹をなす。例えば、図5において、総頻度比率の基準分析値である再平均値で1を除した商が、総頻度比率の補正係数Aとなる。具体的には、0.7445である場合、この0.7445で1を除した商の1.3431が補正係数Aとなる。

【0150】また、段階係数Bは、乖離値が、満点が最適という評価概念ではなく、理想値に近いと最適という評価概念に基づくものであるという性質に鑑み、乖離の程度を分析結果として見たときに、評価数値の整数部分の数で把握しやすいように、算出結果を修正する係数である。本実施形態では、段階係数Bを5に設定している。

【0151】具体的には、本発明で用いる乖離値は、基準となる値に対して、どの程度乖離しているかという程度を表すものである。したがって、例えば0を基準とした場合を考えると、+5であっても、-5であっても、乖離の程度は同じとなる。この場合では、0で表される「ほどほど」のところが最も良いとしたときの、その最

善値からの離れ具合を示す段階を5としたものといえ、評価の内容によっては、3段階、10段階などのほうが感覚的に分かりやすい、または、評価結果を利用しやすい場合があるので、固定値にせず、任意に採り得る係数としている。このため、段階係数Bは、処理手段36による算出結果で得るのではなく、事前に所定の記憶領域に格納される構成を採用する。

【0152】そして、分散係数Cは、乖離値の乖離の程度を示すことになる幅を修正するものである。単に補正係数Aをかけて分析値の桁数を揃えただけでは、数値の変化の幅が普遍化されていないので、乖離値毎の相対比較を行う場合には不適である。このために用いられるのが、式(1)における分散係数Cである。

【0153】本実施形態では、分散係数Cを、基準分析値を求める際に算出可能な標準偏差として、乖離値の普遍化を図っている。上述した基準分析値の規定の仕方に応じて、再平均値を算出した際のデータについて、各分析値とその再平均値を用いて算出した標準偏差を用いてもよい、単なる平均値から求めた標準偏差を用いてもよい。

【0154】ここで、乖離値生成手段20は、生成した乖離値を、パラメータ記憶手段17に設けられた乖離値記憶領域17cに格納する機能も備える。ゆえに、乖離値生成手段20は、詳細には、分析値記憶領域17bと、記憶手段33に設けられた不図示の乖離値算出則記憶領域、および所定の係数記憶領域とを参照し、分析値記憶領域17bから読み出した分析値と、乖離値算出則記憶領域から読み出した前述した式(1)で表される乖離値算出則、および係数記憶領域から読み出した基準分析値に基づき規定された係数と、に基づき、乖離値を生成する手段として機能する。そして、この係数に基づき乖離値を生成することにより、間接的にはあるが基準分析値に基づき乖離値を生成する手段として機能する。

【0155】具体的には、処理手段36が、乖離値を生成する際に、特に、式(1)において、この補正係数Aと分析値の積から1を引くことにより、0を基準とした乖離の程度を示す数値とし、段階係数Bを乗じることにより、この乖離の程度を分析結果として見たときに、認識しやすいように修正し、この分散係数Cで除することにより、算出結果が標準偏差に対する比率という意義を与えて、乖離値の最大値と最小値の幅を相対比較の可能なスケールに修正することにより、乖離値生成手段20が実現され、乖離値が生成される。

【0156】このような乖離値を用いて評価する手法を採用したのは、理解しやすい文章の場合には、分析値が特定の値に集束するという経験的に得られた傾向に基づくものである。なお、参考として、図6に、乖離値生成手段20が、生成した乖離値を格納する乖離値記憶領域17cの一形態として、乖離値を格納するデータベースのデータ構造図を示す。

【0157】また、乖離値の段階係数Bによる乖離値の変動幅の設定は、求められる評価のばらつきに応じた段階評価を踏まえて設定されるものであり、事前に設定されるものであってもよいし、分析途中で、能力評価システム1が入力手段31などを介して何段階評価を行うかを受け付ける構成としてもよい。

【0158】次に、基準指標値記憶手段21は、前述した乖離値生成手段20が生成した複数の乖離値についての比較基準を記憶する手段であり、文章の内容を示すパラメータである乖離値を評価する基準を設定する機能を果たす。

【0159】ここで、詳細は詳述するが、比較基準は、文章の評価に際し、複数の優秀な人材を書き手として得られたパターン指標値を平均して得られた指標値から得られる基準指標値を用いる場合、あるいは、一人の優秀な人材を書き手として得られたパターン指標値から得られる基準指標値を用いる場合などがある。基準指標値記憶手段21は、原則として、これらの基準指標値を事前に基準指標値記憶手段21に記憶する。

【0160】また、本実施形態においては、文章に基づく分析ゆえ、厳密に小数点以下の細かい桁までの一致を判断しても、文章比較として適切な結果を得られないことに鑑み、日本語特有の表現の多様性に起因する数値のばらつき、いうなれば文章の表現の幅からくる「ゆらぎ」を加味して比較できるように乖離値の相対比較により得られるパターンを、比較・評価する基準としている。このため、必要とされる能力を備える場合に、その論述した文章が示す各乖離値の理想的なパターンを意味する指標値を基準指標値として位置づけており、さらには、乖離値をその意義に基づきグループ化した場合のパターンを示すパターン指標値を用いる構成としている。また、相対比較を取り入れることにより、表現技術の優劣に、他の能力が埋没しないように担保することもできる。

【0161】具体的には、表現系に分類される体言率、用言率、付属語率からなる第1ブロック、テーマ系に分類される中心キーワード比率(テーマ設定力)、中心補完単語比率(テーマ説明力)、中心1頻度単語比率(テーマ補足力)からなる第2ブロック、単語構成系に分類される総頻度比率(文章整備力)、出現単語比率(単語構成力)からなる第3ブロック、センテンス系に分類されるセンテンス用言率(行動表現力)、センテンス頻度平均値(センテンス主張力)、センテンス比率(センテンス構成力)からなる第4ブロック、主張系に分類される第1単語頻度比率(主張力)、第2単語頻度比率(主張補足力)、第3単語重複率(主張構成力)、同頻度数比率(論理展開力)、単語回帰値(論理強制力)からなる第5ブロック、に分けて、ブロック毎のパターンを比較する。

【0162】ここで、比較分析による能力評価を行う場

合には、基準指標値記憶手段21に格納される各ブロックの基準指標値も、比較基準となるグループサンプルに応じた数値のばらつきを呈することになる。一方、基準分析による能力評価を行う場合には、各ブロックの基準指標値は、ばらつきのない様な状態を呈することになる。

【0163】なお、本実施形態においては、乖離値の持つ意義と求められる能力との相関関係を考慮して部分的な比較が可能となるようにするため、および比較処理の高速化を図るためにブロック単位での比較手法を採用したが、これに限られるものではなく、ブロック制を採用せず、各々の乖離値を比較する比較手法を採用してもよい。

【0164】乖離値比較手段22は、基準指標値記憶手段21から読み出した基準指標値に基づき、乖離値生成手段20が生成した複数の乖離値同士の相対関係を比較する手段である。ここで、乖離値同士の相対関係とは、上述した乖離値のばらつき状態を比較した関係をいい、ブロック毎に基準指標値を設定した場合にはブロック単位での相対関係となり、乖離値毎に基準指標値を設定した場合には乖離値単位での相対関係となる。

【0165】そして、本実施形態における乖離値比較手段22は、乖離値生成手段20が生成した複数の乖離値に基づき、乖離値相互の相対関係を表すパターン指標値を生成するパターン指標値生成手段22aと、基準指標値記憶手段21から読み出した基準指標値とパターン指標値生成手段22aが生成したパターン指標値と、を比較するパターン比較手段22bと、を有する構成を採用する。前述したように表現の多様性を考慮し、乖離値の相対関係を表すパターン指標値を生成するので、絶対的な数値比較よりも、より文章の持つあいまいさを加味した上で、文章のバランスについての的確な比較を行うことができる。

【0166】詳細には、パターン指標値生成手段22aは、一の乖離値を基準として他の乖離値を相対比較した結果に基づき、乖離値相互の相対関係を表すパターン指標値を生成する手段として機能するものであり、相対比較の一手法として、減算処理を用いる。

【0167】以下、減算処理を用いる具体的態様を説明する。パターン指標値生成手段22aは、相対関係として、一の乖離値を基準として他の乖離値との差を用い、差の値に基づき各乖離値に対応するパターン指標値を生成するという機能を果たすが、差の値の分類パターンは、0以上と0未満の2種類、正の値、0、負の値の3種類、所定のしきい値により正の値と負の値を分け、それに0を加えた5種類とするなどのパターンを採り得る。種類数を増やすにつれて精密な比較が可能となるが、相対差の大きさ（パターンレベル）を加味したい場合には5種類以上を選択することが望ましいといえる。

【0168】例えば、段階係数を5とした場合に5種類

のパターン比較を行う場合には、しきい値を±1として、パターン指標値を+2、+1、0、-1、-2の5種類とする。各ブロックに振り分けられた第1番目の乖離値を基準として、相対比較しパターン化するのであるが、第1番目の乖離値の相対比較した指標値は0となる。そして、第2番目、第3番目と順次乖離値の相対比較を行うが、±1は基準値（第1番目の乖離値）よりその該当する乖離値との差が±1以内であるときに指標値として付与され、±2は基準値（第1番目の乖離値）より、その該当する乖離値との差が1より大きく差がある場合に指標値として付与される。

【0169】したがって、パターン指標値の種類数が2種類の場合は、5の(2-1)乗のパターン指標値の組合せが存在し、パターン指標値の種類数が3種類の場合は、5の(3-1)乗のパターン指標値の組合せが存在し、上述したようにパターン指標値の種類数が5種類の場合は、5の(5-1)乗のパターン指標値の組合せが存在する。

【0170】パターン指標値生成手段22aは、上述したように生成したパターン指標値をパラメータ記憶手段17に設けられたパターン指標値記憶領域17dに記憶する手段としても機能する。このパターン指標値記憶領域17dにおいて、パターン指標値生成手段22aは、文章毎に関連付けてパターン指標値を記憶する。なお、参考として、図7に、パターン指標値生成手段22aが、図6に示した乖離値を用いて生成したパターン指標値乖離値を格納するパターン指標値記憶領域17dの一形態として、パターン指標値を格納するデータベースのデータ構造図を示す。

【0171】そして、このパターン指標値生成手段22aが生成したパターン指標値と、基準指標値記憶手段21から読み出した基準指標値とを、パターン比較手段22bは、比較する機能を果たすが、前述した第1から第5までのブロックについて比較する場合には、全体では、 $25 \times 25 \times 5 \times 25 \times 625 = 48825125$ のパターンの比較を行うことになる。ただし、比較の手法としては、必ずしもすべてのパターン指標値を比較する必要はなく、一般的な文章評価による能力比較を行う際などに、表現系と単語構成系という表現の基礎能力を表すパターン指標値を除外し、 $25 \times 25 \times 625 = 390625$ 通りのパターン比較を行うようにしてもよい。また、適職性を判断する場合などには、第5ブロックの625通りのみを検証するようにしてもよい。これに、後述するキーワードに基づく分析を掛け合わせる分析・比較を行ってもよい。

【0172】このように、乖離値比較手段22は、パターン指標値生成手段22aとパターン比較手段22bとにより、パターンブロック単位の第1番目の乖離値を用いてパターン指標値に変換し、相対比較を行うことにより、パターン認識に擬した比較判断を行い、文章の書き手の能力と基準指標値とを客観的に比較して、その比較

結果をパラメータ記憶手段17に設けられたパターン比較結果記憶領域17eに格納する。そして、一の乖離値を基準として相対関係をみることにより、乖離値の相対関係がより明瞭となる比較対象値の普遍化ができ、また、乖離値の絶対値の大小を比較せず、相対比較する手法を採用することにより、比較パターン数を少なくすることにより、マッチング処理の高速化が図れる。

【0173】ここで、パターン比較手段22bは、比較結果として、パターン指標値の数値を比較し、同じパターンとなっているか否かを比較判断した結果を取得するが、一致する場合には、基準指標値記憶手段21に格納した基準指標値が意味するところの必要とされる能力を備えることになり、一致しない場合には、その能力を備えていないということになる。単に、一致するか否かを結果としてもよいが、一致しなかった部分と一致した部分とを分けた比較結果とすることにより、その人物に備わった能力と欠けた能力とを客観的に示すことができる。そのような形態の比較結果を記憶する構成を採用した場合には、単なる人材登用に際しての能力評価だけでなく、どここの能力を開発すると、最適な人材となるか、というコンピテンシー判断機能を兼ね備えた能力評価システムとしても応用可能となる。

【0174】そして、出力手段23は、処理手段36の指示にしたがい、乖離値比較手段22が行った比較結果を、パターン比較結果記憶領域17eから読み出して出力する。なお、この出力手段23の詳細については、後述する。

【0175】続いて、乖離値比較手段22のように数値的な比較を行う構成に対し、キーワードの意義について比較するための、キーワード抽出手段24、キーワード関連情報記憶手段25、キーワード比較手段26、キーワード使用頻度比較手段27について説明する。これらキーワードの意義について比較する構成を設けたのは、文章構造、表現技術、品詞構成などから判明する能力だけでなく、書き手の想いの強さや、テーマ意識の強さなどをあわせて評価するためである。

【0176】キーワード抽出手段24は、自然文分解手段14が自然文を分解して得られた複数の単語の中から、分析値生成手段18が前述した中心キーワード比率などの分析値の生成に際してキーワードと認定した単語を特定することにより、キーワードを抽出する手段として機能する。

【0177】具体的には、分析値生成手段18が、単語をキーワードとして認定する際には同頻度比率を算出することになるため、このキーワード認定作業にともない、処理手段36は、キーワードとなる単語に抽出用のフラグを立てて、そのフラグと分解単語記憶手段15に格納された単語とを関連付けるなどの処理を事前に行う。そして、そのフラグにしたがって、キーワード抽出手段24は、分解単語記憶手段15からキーワードとな

る単語を抽出し、抽出した単語をキーワード関連情報記憶手段25に記憶する。

【0178】キーワード関連情報記憶手段25は、基準指標値記憶手段21に格納した比較基準に応じた文章に用いられたキーワードを格納する基準キーワード記憶領域25aと、前述したキーワード抽出手段24が抽出した単語を格納するキーワード抽出結果記憶領域25bと、その抽出された単語に関連付けて、後述するキーワード比較手段26やキーワード使用頻度比較手段27が比較した判断結果を記憶する判断結果記憶領域25cと、を備える。

【0179】なお、基準キーワード記憶領域25aに格納される基準キーワードは、比較分析、基準分析ともに、能力があるとされるグループまたは個人の文章を分析し、キーワードと認定された単語が用いられることになる。したがって、分析値生成手段18が、それらの文章を分析する際に、キーワードと認定した単語を基準キーワード記憶領域25aに格納する構成としてもよい。また、この基準キーワード記憶領域25aには、基準キーワードとされた単語の使用頻度を示すパラメータとして、頻度数や同頻度数などのカウント値や分析値などが基準キーワードの付帯情報として関連付けられて格納される。後述するキーワード使用頻度比較手段27が用いることを想定したものである。

【0180】キーワード比較手段26は、キーワード抽出手段24が抽出したキーワードを文章毎に比較し、分析用辞書記憶手段13に設けられた、少なくとも単語の同義語もしくは類義語のいずれかを規定する比較用辞書記憶領域を参照して同義語または類義語の関係にあるか、もしくはキーワードが一致するか、の判断を行う手段として機能する。分解された複数の単語から抽出されたキーワードの同義語や類義語まで比較可能なので、単語の意味内容まで加味した精度の高い文章内容の比較評価を行うことができる。

【0181】ここで、本実施形態における比較用辞書記憶領域とは、図3に示したように分析用辞書記憶手段13に設けられた類義語・同義語辞書記憶領域が該当するが、単に一致するか否かの判断だけでなく、類義語や同義語までも考慮したのは、日本語表現の多様性を鑑みたものである。すなわち、同一の単語を用いなくとも、同趣旨の論述が可能であることから、キーワード比較手段26は、キーワード抽出手段24が抽出したキーワードが、基準キーワード記憶領域25aに格納されたキーワードと一致する場合のみならず、同義または類義のキーワードを用いていると判断した場合にも、高評価とする判断結果を判断結果記憶領域25cに格納する。

【0182】キーワード使用頻度比較手段27は、同義語または類義語の関係にあるか、あるいは、一致するとキーワード比較手段26が判断したキーワードについて、分析値に基づき文章毎の使用頻度を比較する手段と

して機能する。

【0183】このように使用頻度まで評価対象とするのは、同一あるいは同義・類義のキーワードを用いていても、その使用頻度が異なり、例えば、第1キーワードと第3キーワードの関係にあったりする場合には、第1キーワードは異なることになるので、文章中で主として意識したテーマもしくは重要視したポイントが異なる可能性がある。そこで、本実施形態では、キーワード使用頻度比較手段27を設けて使用頻度を比較・分析する構成とした。

【0184】ここで、使用頻度を比較する際に参照する分析値として、キーワード使用頻度比較手段27は、キーワードの文章中で使用された回数としてカウント値取得手段16が取得したカウント値、もしくは、カウント値に基づき分析値生成手段18が生成した他の分析値を用い、これらを比較することにより使用頻度の比較とする。キーワードの文章中で使用された頻度に基づく比較を行うので、文章中における書き手の単語の位置づけまでを含めた評価を実現する。

【0185】なお、キーワード使用頻度比較手段27は、使用頻度を比較した結果を判断結果記憶領域25cに記憶する機能も果たす。

【0186】顕在能力評価値生成手段28は、乖離値生成手段20が生成した複数の乖離値に基づき顕在能力評価値を生成する手段である。顕在能力評価値の例としては、理解する力の評価要素として理解力と把握力が、分析する力の評価要素として分析力、洞察力、および問題発見力が、考える力の評価要素として創造力が、表現する力の評価要素として表現力と説得力が、応用する力の評価要素として応用力、行動力、および対応力が、それぞれ挙げられる。このように、それぞれ異なる観点からなる複数の乖離値を用いて評価要素となる顕在能力が算出される。

【0187】処理手段36が実現する顕在能力評価値の算出方式としては、複数の乖離値の平均値を用いる方式や、乖離値の基準分析値（理想値）と比較する乖離値の差の絶対値をとり、この絶対値に基づいて理想値にどれくらい及ばないかを、理想値を満点とする減点方式で表す方式など、任意の算出方式を採り得る。この算出則も、記憶手段33の所定の領域に設けられた顕在能力評価値算出則記憶領域に格納される。

【0188】そして、顕在能力評価値生成手段28は、他の手段の生成手順と同じく、詳細には、乖離値記憶領域17cと記憶手段33に設けられた不図示の顕在能力評価値算出則記憶領域とを参照し、乖離値記憶領域17cから読み出した乖離値と、顕在能力評価値算出則記憶領域から読み出した顕在能力評価値算出則と、に基づき、顕在能力評価値を生成する手段として機能する。乖離値に基づき顕在能力評価値を生成、出力するので、より具体的な客観評価を実現できる能力評価システムを提

供できる。

【0189】具体的には、本実施形態において、理解する力の評価要素となる理解力の評価は、体言率、テーマ説明力、単語構成力、センテンス構成力、主張補足力という観点から複合的に判断される。体言率、中心補完単語比率、出現単語比率、センテンス比率、第2単語頻度比率、の5つの乖離値の平均をとって数値化する。

【0190】理解する力の評価要素となる把握力の評価は、用言率、テーマ説明力、単語構成力、主張構成力、論理展開力という観点から複合的に判断される。用言比率、中心補完単語比率、出現単語比率、第3単語重複率、同頻度数比率、の5つの乖離値の平均をとって数値化する。

【0191】分析する力の評価要素となる分析力の評価は、体言率、テーマ設定力、テーマ補足力、行動表現力、主張構成力という観点から複合的に判断される。体言率、中心キーワード比率、中心1頻度単語比率、センテンス用言率、第3単語重複率、の5つの乖離値の平均をとって数値化する。

【0192】分析する力の評価要素となる洞察力の評価は、語彙力、テーマ説明力、センテンス主張力、主張力、主張補足力という観点から複合的に判断される。除数詞名詞比率、中心補完単語比率、センテンス頻度平均値、第1単語頻度比率、第2単語頻度比率、の5つの乖離値の平均をとって数値化する。

【0193】分析する力の評価要素となる問題発見力の評価は、体言率、テーマ補足力、単語構成力、行動表現力、主張構成力という観点から複合的に判断される。体言率、中心1頻度単語比率、出現単語比率、センテンス用言率、第3単語重複率、の5つの乖離値の平均をとって数値化する。

【0194】考える力の評価要素となる創造力の評価は、語彙力、用言率、テーマ設定力、主張力、論理強制力、という観点から複合的に判断される。除数詞名詞率、用言比率、中心キーワード比率、第1単語頻度比率、単語回帰値、の5つの乖離値の平均をとって数値化する。

【0195】表現する力の評価要素となる表現力の評価は、語彙力、テーマ補足力、単語構成力、センテンス構成力、主張力、という観点から複合的に判断される。そして、除数詞名詞率、中心1頻度単語比率、出現単語比率、センテンス比率、第1単語頻度比率、の5つの乖離値の平均をとって数値化する。

【0196】表現する力の評価要素となる説得力の評価は、語彙力、テーマ設定力、テーマ補足力、主張力、論理強制力、という観点から複合的に判断される。除数詞名詞率、中心キーワード比率、中心1頻度単語比率、第1単語頻度比率、単語回帰値、の5つの乖離値の平均をとって数値化する。

【0197】応用する力の評価要素となる応用力の評価

は、語彙力、テーマ補足力、単語構力、主張構力、論理展開力という観点から複合的に判断される。除数詞名詞率、中心1頻度単語比率、出現単語比率、第3単語重複率、同頻度数比率、の5つの乖離値の平均をとって数値化する。

【0198】応用する力の評価要素となる行動力の評価は、用言率、センテンス主張力、主張力、論理展開力、論理強制力という観点から複合的に判断される。用言比率、センテンス頻度平均値、第1単語頻度比率、同頻度数比率、単語回帰値、の5つの乖離値の平均をとって数

値化する。

【0199】応用する力の評価要素となる対応力の評価は、テーマ設定力、テーマ説明力、テーマ補足力、行動表現力、主張力という観点から複合的に判断される。中心キーワード比率、中心補完単語比率、中心1頻度単語比率、センテンス用言率、第1単語頻度比率、の5つの乖離値の平均をとって数値化する。

【0200】ただし、評価要素となる力は上述した各力に限られるものではなく、例えば、思考力の評価は、語彙力、テーマ設定力、単語構力、主張力、論理展開力という観点から複合的に判断することができる。この場合、除数詞名詞率、中心キーワード比率、出現単語比率、第1単語頻度比率、同頻度数比率、の5つの乖離値の平均をとって数値化する。

【0201】出力手段23は、乖離値生成手段20が生成した乖離値もしくは顕在能力評価値生成手段28が生成した評価値などの数値結果や、キーワード比較手段26やキーワード使用頻度比較手段が比較した判断結果などを、AGP (Accelerated Graphics Port) やパラレルポート、あるいはUSBなどの情報出力インターフェースを介して外部装置へ出力する手段である。また、出力に際し、出力手段23は、出力先に応じた形態の出力情報の生成を必要に応じて行う。

【0202】例えば、表示手段32に対してパラメータ記憶手段17に格納された情報を出力する場合には、表示手段32においてどのように表示されるかを規定する表示制御情報に乖離値または評価値を組み込んで出力する。印刷手段35に対してパラメータ記憶手段17に格納された情報を出力する場合には、印刷手段35がどのように情報を印刷するかを規定する印刷制御情報に乖離値または評価値を組み込んで出力する。

【0203】機能的には、出力手段23は、乖離値の段階で分析結果の出力を行う場合には、乖離値記憶領域17cから必要な観点の乖離値を読み出し、外部装置である表示手段32や印刷手段35に出力する。評価値による分析結果の出力を行う場合には、出力手段23は、顕在能力評価値記憶領域17fから必要な観点の評価値を読み出し、外部装置である表示手段32や印刷手段35に出力する。また、キーワードの比較結果や使用頻度の比較結果の出力を行う場合には、判断結果記憶領域25

cから必要な比較結果を読み出し、外部装置である表示手段32や印刷手段35に出力する。

【0204】なお、表示手段32や印刷手段35に出力する際に、出力手段23は、乖離値の基準分析値を0から5に修正するなどの出力情報の調整処理を行ってもよい。これは、各乖離値により評価された表3に示した分析項目内容を視覚的に把握しやすいように、基準分析値を全体的に底上げして、すべての評価数値が原則として0以上になるように調整を行うものである。

【0205】なお、要求される出力形態に応じて、そのまま基準分析値を修正せずに乖離値を出力してもよいし、さらには、乖離値などに基づき、グラフ形式や表形式などとなるような出力情報を生成して出力してもよい。なお、表示手段32が出力先である場合には、グラフ表示情報や表形成情報を含む表示制御情報を生成して出力し、印刷手段35が出力先である場合には、グラフ表示情報や表形成情報を含む印刷制御情報を生成して出力する。

【0206】以上説明した各手段が有機的に結合し、機能することにより、自然文により構成された文章の内容という無形のものを対象として、異なる観点を設定した複数のパラメータによる多面的な分析および評価を行うことが可能となり、数値による相対分析および評価が可能になる。そして、客観的に文章内容の評価を行うことができ、さらには、複数の文章について、それぞれ比較結果を出力することにより、文章同士の相対評価も容易に行うことができる。

【0207】特に、文章の内容の優劣を支配する単語の属性に応じた分析区分を設定し、この分析区分に該当する単語の数などに基づく比率を分析値に反映させることにより、文章の評価において重要な使用単語のバランスを数値として評価できる。ひいては、この分析区分の設定を変えることで、目的に応じた能力評価環境を柔軟に提供できる。

【0208】[能力評価処理の手順]次に、能力評価システム1の能力評価処理の手順につき、図8に基づき説明する。図8は、本実施形態にかかる能力評価処理の主要手順を示すフローチャートである。以下に説明する能力評価処理は、前述した乖離値などを用いて評価する数値評価処理と、それに引き続き行われるキーワード判断処理に大別される。

【0209】[数値評価処理]まず、数値評価処理について図8にそって説明する。ここで、本実施形態にかかる数値評価処理を行う能力評価システム1は、数値評価処理を行う際の前提として、上述したように、情報の受付手段と、情報の記憶手段33と、これら各手段の動作を制御する処理手段36と、を備え、記憶手段33が、文章を構成する自然文を単語に分解する基準となる基準単位単語を規定した分析用辞書記憶手段13と、自然文を分解して得られた複数の単語を記憶する分解単語記憶

手段15と、文章に対応する分析値の基準分析値を記憶する基準分析値記憶手段19と、分析値に基づき得られる乖離値についての比較基準を記憶する基準指標値記憶手段21と、を有する構成を採用する。なお、ここでいう受付手段は、入力手段31や通信手段34などが接続されるインターフェース等を想定している。

【0210】そして、この能力評価システム1における数値評価処理および後述するキーワード判断処理を含む能力評価処理は、記憶手段33の所定の領域に格納された能力評価プログラムにしたがって行われる。その能力評価プログラムは、処理手段36に、受付手段を介して自然文からなる文章を受け付けるステップ(S1)と、分析用辞書記憶手段13を参照して、基準単位単語に相当する複数の単語に自然文を分解するとともに、自然文を分解して得られた複数の単語を分解単語記憶手段15に記憶するステップ(S2)と、分解単語記憶手段15が記憶した複数の単語を参照し、少なくともその単語にかかわる数をカウント値として取得するステップ(S3)と、カウント値に基づき複数の分析値を生成するステップ(S4)と、生成した分析値と基準分析値記憶手段19が記憶する基準分析値とに基づき、基準分析値に対する乖離の程度を示す複数の乖離値を生成するステップ(S5)と、基準指標値記憶手段21から読み出した基準指標値に基づき、生成した複数の乖離値同士の相対関係を比較するステップ(S6)と、相対関係についての比較結果を出力するステップ(S9)と、を実行させる。

【0211】これらの各ステップを実行させることにより、客観的に文章内容の評価を行うことができ、ひいては、複数の文章について、それぞれ比較結果を出力することにより、文章同士の相対評価も容易に行うことができる能力評価環境を提供するものである。

【0212】より詳細には、図8に示したように、まず、処理手段36に自然文を受け付けるステップ(S1)を実行させる。このとき、入力手段31から文章が直接的に入力される場合には、入力手段31が接続されたPS/2(Personal System/2)ポートやUSB(Universal Serial Bus)などが受付手段となり、処理手段36は、これらを介して自然文からなる文章を受け付け、受け付けた文章を文章記憶手段12の所定の領域に格納する。

【0213】また、外部からの通信手段34から文章を電子情報として受け付ける場合には、通信手段34が接続されたRS232C(Recommended Standard 232 version C)ポートやNIC(Network Interface Card)、USBなどを受付手段となり、処理手段36は、これらを介して自然文からなる文章を受け付け、受け付けた文章を文章記憶手段12の所定の領域に格納する。

【0214】加えて、可搬な記憶媒体もしくは内部の記憶手段33に格納された文章情報を用いる場合には、P

CIバス(Peripheral Components Interconnect bus)などが受付手段となり、処理手段36は、文章情報を分析対象として受け付けるとともに、受け付けた文章情報を文章記憶手段12の所定の領域に格納する。

【0215】そして、上述したように、受付手段は文章の受付の形態により異なるが、処理手段36に受付手段を介して文章を受け付け、受け付けた文章を文章記憶手段12の所定の領域に格納することを実行させることにより、前述した文章受付手段11が実現され、機能する。

【0216】なお、本実施形態では、自然文を単語に分解する次のステップ(S2)に進む前に、処理手段36に、文章記憶手段12に格納された文章の全体の文字数をカウントさせるステップ、分析用辞書記憶手段13の句点識別辞書記憶領域に格納されている句点を一致判断基準とした文章のセンテンス単位への分解ステップ、およびセンテンス数のカウントステップなど、単語に基づくカウント値以外の分析値を生成するのに必要なカウント値の取得ステップ等を実行させる。

【0217】その上で、さらに、分析対象となる文章を構成する自然文について、処理手段36に、分析用辞書記憶手段13の表記辞書記憶領域と削除記号辞書記憶領域とを参照し、記号交換や記号削除などを施して、単語への分解に適した自然文に加工するステップを実行させた後、単語分解ステップ(S2)へと移行する。

【0218】これらの前処理に続いて、図8に示したように、処理手段36に、分析用辞書記憶手段13の体言辞書記憶領域、用言辞書記憶領域、接続詞辞書記憶領域、助詞辞書記憶領域、助動詞辞書記憶領域、および活用辞書記憶領域などを参照して、基準単位単語に相当する複数の単語に自然文を分解するとともに、自然文を分解して得られた複数の単語を分解単語記憶手段15に記憶するステップ(S2)を実行させる。

【0219】具体的には、処理手段36は、1センテンス単位で自然文の頭から単語に分解し、分解された順番に記憶手段33の所定の領域に一時的に格納する。また、処理手段36は、分解に際し活用辞書記憶領域を参照するが、この参照により、用言を終止形に置き換えるなどの処理を行う。

【0220】ここで、本実施形態では、前述したように、分析用辞書記憶手段13が、基準単位単語の属性に応じた分析区分を基準単位単語に関連付けて記憶していることが前提となる。そして、詳細には、処理手段36に、複数の単語に自然文を分解する際に、単語に分析区分を関連付けるように、分解対象となる単語と基準単位単語の一致を判断した上で、基準単位単語に付与されている分析区分をそのまま分解した単語に付与する処理を実行させ、その単語に付与した分析区分を分解した単語とともに分解単語記憶手段15に記憶することにより、分析区分と関連付けられた単語を分解単語記憶手段15

に記憶するステップ(S2)を実行させる。なお、分析区分の付与は、表1および表2に示した分析区分コードを付与する方式を採用する。

【0221】このように、処理手段36に、分析用辞書を参照しながら自然文を単語に分解し、分解した単語を分解単語記憶手段15の所定の領域に格納することを実行させることにより、前述した自然文分解手段14が実現され、機能する。

【0222】そして、上述した分解ステップ(S2)の後、図8に示したように、処理手段36に、分解単語記憶手段15が記憶した結果を参照し、少なくとも分解した単語の数をカウント値として取得するとともに、取得したカウント値をカウント値記憶領域17aの所定の領域に格納させるステップ(S3)と、カウント値に基づき複数の分析値を生成するステップ(S4)と、を実行させるものであるが、これらのステップは、生成対象となる分析値の生成過程において、必要に応じて繰り返されることになる。以下、具体的に各ステップについて説明する。

【0223】まず、カウント値を取得するステップ(S3)として、処理手段36は、分解単語記憶手段15に格納されている自然文を分解して得られた複数の単語を参照して、少なくとも単語にかかわる数を分析対象とした文章についてのカウント値として、例えば、前述した文章に固有の識別情報などに対応させて取得する。そして、処理手段36は、取得したカウント値をカウント値記憶領域17aの所定の領域に格納する。この際に取得、格納されるカウント値としては、前述した総頻度数や体言数、同頻度数などが挙げられる。

【0224】このように、処理手段36に所定の領域に格納されている情報に基づき、所定の情報単位をカウントさせ、カウント結果をカウント値として取得、格納することを実行させることにより、前述したカウント値取得手段16が実現され、機能する。

【0225】続いて、カウント値に基づき複数の分析値を生成するステップ(S4)として、処理手段36は、カウント値記憶領域17aに格納されているカウント値を読み出すとともに、前述した表3に示した分析値算出則を記憶手段33に設けられた所定の分析値算出則記憶領域から読み出し、読み出した分析値算出則に基づきカウント値を用いて分析値を生成する。分析値の生成に際しては、図4に示したように、複数の分析対象について複数種類の分析値を算出することになるので、結果として、処理手段36は、カウント値に基づき複数の分析値を複数種類生成させることになる。

【0226】より詳細には、処理手段36は、分解単語記憶手段15に記憶された分解された単語と、その単語と分析区分との関連付けを参照し、分析区分に属する単語をカウントしたカウント値に基づく複数の分析値を生成する。この際に生成される分析値として、センテンス

用言率などが挙げられる。

【0227】また、必要に応じて、処理手段36が生成したカウント値から生成された二次的な数値もしくは他の分析値を用いて、処理手段36は、前述した分析値算出則記憶領域などから読み出した分析値算出則にしたがって、分析値を生成する。この際に生成される分析値としては、体言率や用言率などが挙げられる。そして、処理手段36は、生成した分析値を分析値記憶領域17bの所定の領域に格納する。

10 【0228】ここで、分析値生成ステップ(S4)からカウント値取得ステップ(S3)へと戻る態様として、キーワード認定処理が挙げられる。具体的には、処理手段36は、キーワード比率を生成するために、カウント値として得た同頻度数に基づき、単語の同頻度件数などを取得し、同頻度比率を取得する。このとき、同頻度比率が0.7以下の単語がキーワードと認定されるが、このときの認定された単語数のカウントのために、再度、カウント値取得ステップ(S3)へと戻る。また、同じく、キーワードと認定された単語群の中に含まれる感想

20 単語の数を取得するためにも、処理手段36は、カウント処理を行う。

【0229】この後、キーワード数などのカウント値に基づき、処理手段36は、中心キーワード数などの分析値を生成する。

30 【0230】ただし、厳密には、記憶手段33の所定の領域に格納された、キーワードに認定された単語に関連するデータを参照するため、前述した自然文分解手段14が分解した結果に基づくカウントを行うステップとは、カウント対象が異なるカウント値取得ステップと言える。

【0231】また、処理手段36は、上述したキーワード認定処理にともない、キーワードとして認定した単語について、キーワード識別コードなどのフラグを立て、キーワードとして認定した単語自体を抽出できるようにする処理を施して、再度、分解単語記憶手段15に格納する。また、処理手段36は、その単語の頻度数もしくは同頻度比率なども、その単語に関連付けて記憶手段33の所定の領域に記憶する。

40 【0232】このように、処理手段36にカウント値に基づく分析値の生成を実行させることにより、前述した分析値生成手段18が実現され、機能することになる。

【0233】次に、分析値から算出した基準値に対する乖離の程度を示す乖離値を生成するステップ(S5)として、処理手段36は、分析値記憶領域17bから分析値の種類毎に複数の分析値を読み出し、読み出した分析値と基準分析値記憶手段19が記憶する基準分析値とに基づき、記憶手段33に設けられた所定の乖離値算出則記憶領域に格納された乖離値算出則を用いて、基準分析値に対する乖離の程度を示す複数の乖離値を生成する。

50 【0234】具体的には、処理手段36に、前述した式

(1)を記憶手段33の所定の領域から読み出し、分析値の種類毎に対応する基準値から得られた補正係数などの各係数を基準分析値記憶手段19から読み出し、読み出した係数と式(1)に基づいて各分析値について乖離値を生成するステップを実行させる。そして、処理手段36は、生成した乖離値を乖離値記憶領域17cに記憶する。

【0235】上述したように、処理手段36に、前述した式(1)のような所定の算出則に基づき、分析値を用いて乖離値を生成させるとともに、生成した乖離値を乖離値記憶領域17cに格納させることを実行させることにより、前述した乖離値生成手段20が実現され、機能することになる。そして、その機能が発揮されることにより、複数種類の分析値を用いて複数の乖離値を得るので、多面的な観点から評価できる要素を含んだ乖離値を得ることができ、より実態を反映した文章内容の評価を実現できる。

【0236】続いて、乖離値同士の相対関係を比較するステップ(S6)として、処理手段36は、基準指標値記憶手段21から読み出した基準指標値に基づき、生成した複数の乖離値同士の相対関係を比較し、比較した比較結果をパターン比較結果記憶領域17eに記憶する。詳細には、処理手段36に、乖離値同士の相対関係を比較するステップとして、複数の乖離値に基づき、乖離値相互の相対関係を表すパターン指標値を生成するステップと、生成したパターン指標値と、基準指標値記憶手段21から読み出した基準指標値とを比較するとともに、比較したパターン比較結果を記憶手段33に設けられたパターン比較結果記憶領域に記憶するステップと、を実行させる(S6)。

【0237】特に、本実施形態では、文章毎の乖離値の相対関係を比較するに際し、処理手段36に、乖離値相互の相対関係を表すパターン指標値を、一の乖離値を基準として他の乖離値を相対比較した結果に基づき生成するステップを実行させる手法を採用する。具体的には、処理手段36に、相対関係として、一の乖離値を基準として他の乖離値との差を算出し、その差の値に基づき各乖離値に対応するパターン指標値を生成するステップを実行させる(S6)。

【0238】本実施形態でのパターン指標値生成においては、乖離値の厳密な差の数値結果を用いることなく、全体的な分析パターンの比較という思想を実現するために、乖離値の数値ではなく、乖離値同士の相対関係を表すパターン指標値に置き換える処理をおこなう。この処理により、より文章の持つあいまいさを加味した上で、文章のバランスについての的確な比較を行うことができる。具体的には、パターン全体の比較を行う場合には、好適なパターン指標値として、-1、0、+1という3種類のパターン指標値を用い、パターンの特性をより反映させた比較を行う場合には、パターンレベルを考慮し

たパターン指標値として、前述したように乖離値の生成における段階係数を5にしたことから、±1をしきい値として用い、-2、-1、0、1、2という5種類のパターン指標値を用いる。

【0239】ただし、前述したように、処理手段36は、すべての乖離値について相対関係を取得する必要はなく、能力判定に必要な乖離値相互の関係をパターン指標値として生成・取得すればよい。

【0240】また、パターン指標値の比較するステップとして、処理手段36は、基準指標値記憶手段21から、能力評価の理想パターンである基準指標値を読み出し、前述した文章毎に生成されたパターン指標値との比較を行う。処理手段36は、比較処理として、パターン指標値の数値について一致、不一致を判断し、判断したパターン指標値毎の一致、不一致という比較結果をパターン比較結果記憶領域17eに記憶する。

【0241】このように、処理手段36に、乖離値相互の相対関係を表すパターン指標値を、一の乖離値を基準として他の乖離値を相対比較した結果に基づき生成することを実行させることにより、前述したパターン指標値生成手段22aが実現され、機能する。また、処理手段36に、生成したパターン指標値と、基準指標値記憶手段21から読み出した基準指標値とを比較することを実行させることにより、前述したパターン比較手段22bが実現され、機能する。

【0242】なお、顕在能力評価値を生成する場合には、処理手段36は、乖離値記憶領域17cから乖離値を読み出し、さらには、前述したように、所定の領域に格納された評価値の算出則にしたがい、読み出した複数の乖離値に基づき顕在能力評価値を生成し、生成した評価値を顕在能力評価値記憶領域17fに格納する。このように、処理手段36に、乖離値に基づき評価値を生成させ、生成させた評価値を格納させることにより、前述した顕在能力評価値生成手段28が実現され、機能する。そして、その機能が発揮されることにより、複数の乖離値に基づき、評価値を生成するので、異なる観点から多面的に分析したパラメータから構成される信頼度の高い評価値を得ることができ、ひいては、文章内容の数値評価環境として高度なシステムを提供することができる。

【0243】上述した数値評価処理の後、処理手段36は、キーワード判断処理を行うか否かの指示を受け付ける判断ステップを実行する(S7)。処理手段36は、入力手段31などを介して指示を受け付ける。ここで、キーワードを比較してその意義についての認識までを評価する旨の指示を受け付けた場合には、キーワードの内容に基づき評価する一連のステップ(S8)へと移行する。一方、キーワード判断処理を行わない旨の指示を受け付けた場合には、出力ステップ(S9)へと移行する。

【0244】なお、この判断ステップ(S7)は、例えば、処理手段36が生成した表示制御情報に基づき、表示手段32に形成した質問画面(指示受付インターフェース)を介して、入力手段31から指示命令を受け付けるなどの処理により実現される構成としてもよい。

【0245】[キーワード判断処理] 続いて、キーワード判断処理について説明する。ここで、前述したように、キーワード判断処理は、記憶手段33の所定の領域に格納された能力評価プログラムにしたがって行われるものであり、処理手段36に、キーワードの内容に基づき評価する一連のステップ(S8)として、自然文を分解して得られた複数の単語の中から、キーワードを抽出するステップと、キーワードを文章毎に比較し、前記比較用辞書記憶領域を参照して同義語または類義語の関係にあるか、もしくはキーワードが一致するか、の判断を行うステップと、判断の結果を出力するステップと、を実行させる。これらのステップの実行により、分解された複数の単語から抽出されたキーワードの同義語や類義語まで比較可能なので、単語の意味内容まで加味した精度の高い文章内容の評価を行うことができる。

【0246】より詳細には、まず、処理手段36に自然文を分解して得られた複数の単語の中から、キーワードを抽出するステップ(S8)を実行させる。ここで、処理手段36は、分解単語記憶手段15を参照し、前述したキーワード識別コードなどのフラグを抽出条件として、キーワードと認定された単語を抽出し、抽出した単語をキーワードとして、キーワード抽出結果記憶領域25bに記憶する。ただし、識別コードなどを用いず、頻度数などのカウント値や同頻度比率などの分析値を抽出条件としてもよい。そして、この処理を実行させることにより、前述したキーワード抽出手段24が実現され、機能する。

【0247】抽出処理に続く比較ステップとして、処理手段36に、キーワードを文章毎に比較し、前記比較用辞書記憶領域を参照して同義語または類義語の関係にあるか、もしくはキーワードが一致するか、の判断を行うステップを実行させる(S8)。処理手段36は、キーワード抽出結果記憶領域25bから分析対象となる文章のキーワードを読み出し、読み出したキーワードを、基準キーワード記憶領域25aに格納されている基準キーワードと比較する。この直接比較で一致する場合は、処理手段36は、一致した旨を、判断結果記憶領域25cに判断結果として記憶する。直接比較で一致しなかった場合には、処理手段36は、分析用辞書記憶手段13に備えられた類義語・同義語辞書記憶領域を参照し、基準キーワードの同義語・類義語を特定し、特定した同義語・類義語と分析対象文章のキーワードとを比較する。この二次比較で一致する場合は、処理手段36は、一致した旨を、判断結果記憶領域25cに判断結果として記憶する。また、直接比較、二次比較ともに不一致であった

場合には、その旨を、判断結果記憶領域25cに判断結果として記憶する。

【0248】このように、処理手段36に、必要に応じて、比較用辞書としての同義語・類義語辞書を参照してキーワードを比較させ、一致・不一致を判断させる処理を実行させることにより、前述したキーワード比較手段26が実現され、機能する。

【0249】そして、上述した比較ステップの後、キーワードの意識の程度までを考慮した能力評価を行う場合には、処理手段36に、同義語または類義語の関係にあるか、あるいは、一致すると前記キーワード比較手段が判断したキーワードについて、分析値に基づき文章毎の使用頻度を比較するステップ(S8)を実行させる。

【0250】具体的には、処理手段36に、前述した文章に固有の識別情報を検索キーとして、カウント値記憶領域17aや分析値記憶領域17bを参照して、キーワード比較対象とした文章に対応する使用頻度として、キーワードの文章中で使用された回数として取得した頻度数などのカウント値、またはカウント値に基づき生成した同頻度数比率などの分析値を読み出し、基準キーワード記憶領域25aから基準キーワードの付帯情報として読み出した基準キーワードの頻度数や同頻度数比率などと比較し、比較した結果を判断結果記憶領域25cに記憶する。

【0251】この使用頻度の比較ステップにおいては、厳密な数値比較をして一致・不一致を判断させてもよいが、より好ましくは、処理手段36に、前述したパターン指標値のように、一致するか、より大きいのか、より小さいか、などの程度の表す情報を判断させ、その旨を判断結果として判断結果記憶領域25cに格納させるとよい。

【0252】そして、使用頻度比較処理においても、上述したように、処理手段36にキーワードの使用頻度を比較させる処理を実行させることにより、前述したキーワード使用頻度比較手段27が実現され、機能する。

【0253】[出力処理] 最後に、出力ステップ(S9)について説明する。図8に示したように、この出力ステップ(S9)へ移行してくる場合には、乖離値や顕在能力評価値の出力を目的とする場合のみならず、パターン指標値の比較判断結果、キーワードの意義・意味の比較判断結果やキーワード使用頻度の比較判断結果などの出力を目的とする場合がある。

【0254】乖離値や顕在能力評価値を出力するように、出力ステップ(S9)へ移行した場合には、処理手段36は、対象となる数値をパラメータ記憶手段17の所定の記憶領域から読み出し、出力先に応じた情報に乖離値や顕在能力評価値を含める加工などを行った上で、出力要求に応じた出力を行う。

【0255】同様に、比較判断結果を出力するように、出力ステップ(S9)へ移行した場合には、処理手段3

6は、パラメータ記憶手段17のパターン比較結果記憶領域17e、キーワード関連情報記憶手段25の判断結果記憶領域25cから比較結果や判断結果を読み出し、それらの結果に基づき出力先に応じた情報に加工などを行った上で、出力要求に応じた出力を行う。

【0256】このように、処理手段36に、所定の出力先へ要求された情報を出力させることにより、前述した出力手段23が実現され、機能することになる。

【0257】なお、参考までに、パターン比較結果と顕在能力評価値との双方に基づき、情報を加工した出力例を図9に示す。図9の出力結果は、表形式で文章毎に結果を一覧できるように表示手段32に出力した場合の画面構造図である。図9に示したように、職種Aおよび職種Bについてパターン指標値の比較を行った場合に、5種類のパターン指標値（パターンレベル）が一致したという結果を「○」で、5種類では一致しなくても3種類のパターン指標値が一致したという結果を「△」で、どちらも一致しなかったという結果を「×」で表すという情報加工を施し、さらに、各文章に基づく分析結果として顕在能力評価値をあわせて出力したときは、このような画面構造を表示するための表示制御情報を出力手段23が出力する。

【0258】以上説明した能力評価処理により、能力評価システム1は、分析対象となる文章を分析、評価するものであり、その文章が書かれた目的などに応じて、客観的な評価を行うことができる、数値化されたパラメータを用いて評価を可能とする能力評価システムおよび能力評価プログラムの提供を実現するものである。

【0259】〔その他の処理〕本実施形態の能力評価システム1を用いて基準分析を行う際に、能力評価に先立って、前述したように優秀な人材の文章に基づき基準分析値を求める場合には、処理手段36に、図8に示した分析値生成までの処理（S1～S4）までを実行させた後、分析値記憶領域17bから、優秀な人材の文章についての分析値の種類毎に複数の分析値を読み出すとともに、記憶手段33に設けられた所定の基準分析値算出則記憶領域に格納された算出則に基づき、読み出した分析値を用いて生成する。

【0260】具体的には、基準分析値算出則記憶領域から所定の算出則を読み出し、処理手段36に、優秀な人材の複数の文章毎に生成された複数の分析値を用いて標準偏差を算出し、算出した標準偏差を用いて再度分析値をフィルタリングして抽出した分析値集団を平均し、分析値の基準となる基準分析値を生成し、生成した基準分析値を基準分析値記憶手段19に格納する処理を実行させる。また、処理手段36に、認定したキーワードを基準キーワード記憶領域25aに格納させるとともに、そのキーワードに関連するカウント値や分析値などを付帯情報としてキーワードに関連付けて格納する処理を実行させる。

【0261】このとき、処理手段36は、前述したように、（平均値）±（標準偏差）の範囲に該当しない分析値については除外する処理をおこない、分析値の再平均値に基づき標準偏差を再度求める処理を行う。

【0262】このように、処理手段36に、基準分析値算出則記憶領域から読み出した平均を求める算出則に基づき、分析値を用いて基準分析値を生成させるとともに、基準分析値記憶手段19に生成した基準分析値を格納させることを実行させることにより、前述した不図示の基準値生成手段が実現され、機能することになる。そして、ひいては、一連の処理により、前述した絶対分析や基準分析などを行うことができる好適な能力評価システム1の提供を実現できる。

【0263】〔第2実施形態〕

【0264】本発明の他の実施形態を図10に基づき説明する。ここで、上述した第1実施形態と同一の構成部分については、同一の符号を付して重複説明を省略する。図10は、ネットワークに対応した能力評価システム2を実現するための概略的なシステム構成を示したシステム構成図である。

【0265】図10に示したように、本実施形態では、能力評価システム2は、ネットワーク4に接続されたサーバ装置として提供されるものであり、能力評価システム2には、ネットワーク4を介して端末装置3が接続される。なお、ネットワーク4は、インターネットや専用回線によるネットワーク、企業内LAN（Local Area Network）などが想定される。また、端末装置3は、パーソナルコンピュータなどの情報端末が想定される。

【0266】能力評価システム2は、記憶手段33と、通信手段37を接続するための不図示のインターフェースと、記憶手段33とそのインターフェースの動作を制御する処理手段36を備える。ここで、通信手段37は、外部ネットワーク4との情報の授受を可能とする手段であり、例えば、ルータなどが挙げられる。この通信手段37に応じたインターフェースを、能力評価システム2が備える。このインターフェースとしては、例えば、LANカードなどのNICが挙げられる。

【0267】ただし、通信手段37が直接にネットワーク4に接続されることは必須とされず、能力評価システム2が通信手段37を介したLAN経由で、プロキシサーバなどの制御の下、ネットワーク4に接続する構成を採用してもよい。

【0268】端末装置3は、入力手段31、表示手段32、通信手段34、印刷手段35、記憶手段38、および各手段の動作を制御する処理手段39を備える。ここで、記憶手段38は、前述した記憶手段33と同様、情報を展開、格納する領域を提供するための手段であり、例えば、メモリ、HDDなどが挙げられる。また、処理手段39は、前述した処理手段36と同様、各手段の管理を行うCPUなどの演算処理手段である。

【0269】詳細には、本実施形態における能力評価システム2は、前述した第1実施形態と同様、図1に示したシステム構成を採用しており、自然文からなる文章を受け付ける文章受付手段11と、当該自然文を単語に分解する基準となる基準単位単語を規定した分析用辞書記憶手段13と、分析用辞書記憶手段13を参照して、前記基準単位単語に相当する複数の単語に前記自然文を分解する自然文分解手段14と、自然文分解手段14が自然文を分解して得られた複数の単語を記憶する分解単語記憶手段15と、分解単語記憶手段15が記憶した複数の単語を参照し、少なくとも当該単語にかかわる数を前記文章についてのカウント値として取得するカウント値取得手段16と、カウント値取得手段16が取得したカウント値に基づき複数の分析値を生成する分析値生成手段18と、当該分析値の基準分析値を記憶する基準分析値記憶手段19と、基準分析値記憶手段19から読み出した基準分析値と、生成された前記分析値と、に基づき当該基準分析値に対する当該分析値の乖離の程度を示す複数の乖離値を生成する乖離値生成手段20と、当該複数の乖離値についての比較基準を記憶する基準指標値記憶手段21と、基準指標値記憶手段21から読み出した基準指標値に基づき、乖離値生成手段20が生成した複数の前記乖離値同士の相対関係を比較する乖離値比較手段22と、乖離値比較手段22が行った比較結果を出力する出力手段23と、を主たる構成として備える。

【0270】そして、さらに、図1に示した構成と同じく、能力評価システム2は、自然文分解手段14が自然文を分解して得られた複数の単語の中から、キーワードを抽出するキーワード抽出手段24と、キーワード抽出手段24が抽出したキーワードを文章毎に比較し、分析用辞書記憶手段13が備える、少なくとも単語の同義語もしくは類義語のいずれかを規定する前記比較用辞書記憶領域を参照して同義語または類義語の関係にあるか、もしくはキーワードが一致するか、の判断を行うキーワード比較手段26と、同義語または類義語の関係にあるか、あるいは、一致するとキーワード比較手段26が判断したキーワードについて、前記分析値に基づき文章毎の使用頻度を比較するキーワード使用頻度比較手段27と、乖離値生成手段20が生成した複数の乖離値に基づき複数の頭在能力評価値を生成する頭在能力評価値生成手段28と、を備える。

【0271】加えて、能力評価システム2における能力評価処理も、原則として図8に示した手順と変わりはない。ただし、本実施形態における能力評価システム2は、ネットワークに対応するものであるため、第1実施形態における情報の受付手段に相当する受付手段は、上述した通信手段37が接続された不図示のインターフェースとなる。

【0272】ゆえに、能力評価システム2における自然文の受付は、詳細には以下ようになる。まず、端末装

置3の処理手段39が、入力手段31を介して、もしくは通信手段34を介して自然文からなる文章を取得し、取得した文章を記憶手段38の所定の領域に格納する。そして、端末装置3の処理手段39は、能力評価を行う対象としての文章を、通信手段34およびネットワーク4を介して、能力評価システム2側へ送信する。これを受けて、能力評価システム2の処理手段36は、通信手段37と不図示のインターフェースを介して、分析対象となる文章を受け付ける。

【0273】処理手段36に、このインターフェースを介した自然文の受け付けを実行させることにより、本実施形態における文章受付手段11が実現され、機能することになる。

【0274】この後、能力評価システム2が、自然文を分解し、カウント値を取得し、分析値、乖離値、パターン指標値などを生成する処理や、キーワードの内容について判断する処理を行うのは、第1実施形態と同様である。

【0275】そして、能力評価システム2が、生成した乖離値や判断結果の出力を行う出力手段23を備える点も同様であるが、本実施形態では、出力手段23の出力先が分析対象を送信してきた端末装置3などのネットワーク端末が想定される。詳細には、能力評価システム2の処理手段36は、記憶手段33に設けられたパラメータ記憶手段17から乖離値、頭在能力評価値もしくはパターン比較結果を読み出し、あるいは、キーワード関連情報記憶手段25から判断結果を読み出し、端末装置3側の表示手段32または印刷手段35など、所定の出力先のハードウェア仕様に応じた情報加工を加えた上で、上述したインターフェースおよび通信手段37を介して端末装置3側へ送信する。

【0276】処理手段36に、このインターフェースを介した乖離値または評価値の送信処理を実行させることにより、本実施形態における出力手段23が実現され、機能することになる。

【0277】

【発明の効果】本発明は、以上のように構成され機能するので、これによると、請求項1に記載の発明では、分解された複数の単語に基づき得られる、所定の基準分析値に対する乖離値を用いて乖離値同士の相対関係を比較するので、客観的に文章内容の評価を行うことができる。ひいては、複数の文章について、それぞれ比較結果を出力することにより、文章同士の相対評価も容易に行うことができる。

【0278】また、請求項2に記載の発明では、乖離値の相対関係を表すパターン指標値を生成するので（絶対的な数値比較よりも）、より文章の持つあいまいさを加味した上で、文章のバランスについての的確な比較を行うことができる。

【0279】また、請求項3に記載の発明では、一の乖

離値を基準として相対関係をみるので、乖離値の相対関係がより明瞭となる普遍化ができ、また、比較パターン数を少なくすることにより、マッチング処理の高速化が図れる。

【0280】また、請求項4に記載の発明では、分解された複数の単語から抽出されたキーワードの同義語や類義語まで比較可能なので、単語の意味内容まで加味した精度の高い文章内容の比較評価を行うことができる。

【0281】また、請求項5に記載の発明では、キーワードの文章中で使用された頻度に基づく比較を行うので、文章における書き手の単語の位置づけまでを含めた評価を実現する。

【0282】また、請求項6に記載の発明では、乖離値に基づき顕在能力評価値を生成、出力するので、より具体的な客観評価を実現できる能力評価システムを提供できる。

【0283】また、請求項7に記載の発明では、分解された複数の単語に基づき得られる、所定の基準分析値に対する乖離値を用いて乖離値同士の相対関係を比較するので、客観的に文章内容の評価を行うことができる。ひいては、複数の文章について、それぞれ比較結果を出力することにより、文章同士の相対評価も容易に行うことができる。

【0284】また、請求項8に記載の発明では、乖離値の相対関係を表すパターン指標値を生成するので、より文章の持つあいまいさを加味した上で、文章のバランスについての確かな比較を行うことができる。

【0285】また、請求項9に記載の発明では、一の乖離値を基準として相対関係をみるので、乖離値の相対関係がより明瞭となる普遍化ができ、また、比較パターン数を少なくすることにより、マッチング処理の高速化が図れる。

【0286】また、請求項10に記載の発明では、分解された複数の単語から抽出されたキーワードの同義語や類義語まで比較可能なので、単語の意味内容まで加味した精度の高い文章内容の評価を行うことができる。

【0287】また、請求項11に記載の発明では、キーワードの文章中で使用された頻度に基づく比較を行うので、文章における書き手の単語の位置づけまでを含めた評価を実現する。

【0288】また、請求項12に記載の発明では、乖離値に基づき顕在能力評価値を生成、出力するので、より具体的な客観評価を実現できる能力評価プログラムを提供できる。

【図面の簡単な説明】

【図1】本発明に係る能力評価システムのシステム構成を示す機能ブロック図である。

【図2】本発明に係る能力評価システムを実現するための概略的なシステム構成を示すシステム構成図である。

【図3】分析用辞書記憶手段のデータ構造の一例を示す

模式的構造図である。

【図4】カウント値記憶領域へのカウント値の記憶形態の一例を示すデータ構造図である。

【図5】基準分析値記憶手段への基準分析値の記憶形態の一例を示すデータ構造図である。

【図6】乖離値記憶領域への乖離値の記憶形態の一例を示すデータ構造図である。

【図7】パターン指標値記憶領域へのパターン指標値の記憶形態の一例を示すデータ構造図である。

10 【図8】本実施形態における能力評価処理を示すフローチャートである。

【図9】表示手段の表示領域に形成される画面に形成される出力結果の一例を示す画面構成図である。

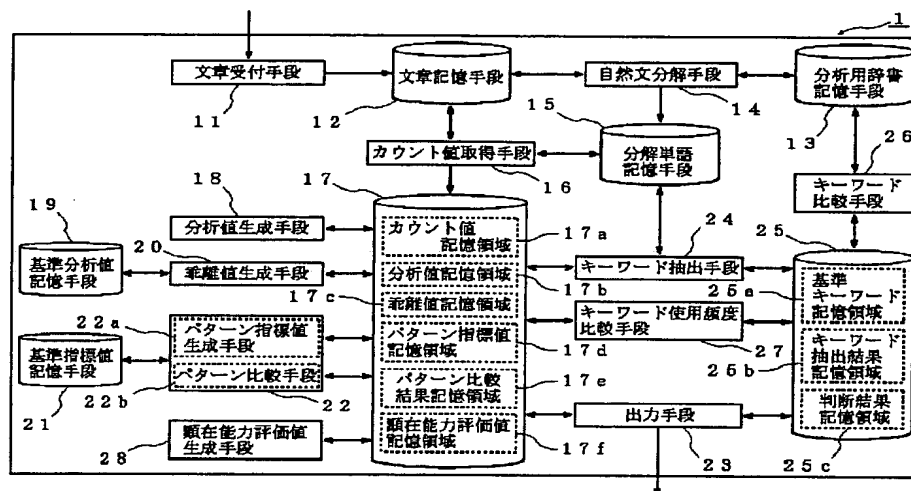
【図10】本発明に係る他の能力評価システムを実現するための概略的なシステム構成を示すシステム構成図である。

【符号の説明】

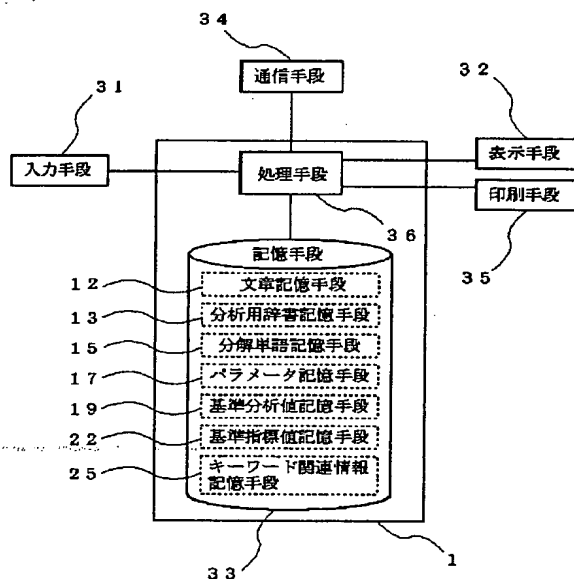
- | | |
|---------|---------------|
| 1 | 能力評価システム |
| 2 | 能力評価システム |
| 20 3 | 端末装置 |
| 4 | ネットワーク |
| 11 | 文章受付手段 |
| 12 | 文章記憶手段 |
| 13 | 分析用辞書記憶手段 |
| 14 | 自然文分解手段 |
| 15 | 分解単語記憶手段 |
| 16 | カウント値取得手段 |
| 17 | パラメータ記憶手段 |
| 17 a | カウント値記憶領域 |
| 30 17 b | 分析値記憶領域 |
| 17 c | 乖離値記憶領域 |
| 17 d | パターン指標値記憶領域 |
| 17 e | パターン比較結果記憶領域 |
| 17 f | 顕在能力評価値記憶領域 |
| 18 | 分析値生成手段 |
| 19 | 基準分析値記憶手段 |
| 20 | 乖離値生成手段 |
| 21 | 基準指標値記憶手段 |
| 22 | 乖離値比較手段 |
| 40 22 a | パターン指標値生成手段 |
| 22 b | パターン比較手段 |
| 23 | 出力手段 |
| 24 | キーワード抽出手段 |
| 25 | キーワード関連情報記憶手段 |
| 25 a | 基準キーワード記憶領域 |
| 25 b | キーワード抽出結果記憶領域 |
| 25 c | 判断結果記憶領域 |
| 26 | キーワード比較手段 |
| 27 | キーワード使用頻度比較手段 |
| 50 28 | 顕在能力評価値生成手段 |

- | | |
|-------|------|
| * 3 6 | 処理手段 |
| 3 7 | 通信手段 |
| 3 8 | 記憶手段 |
| 3 9 | 処理手段 |

【図 1】



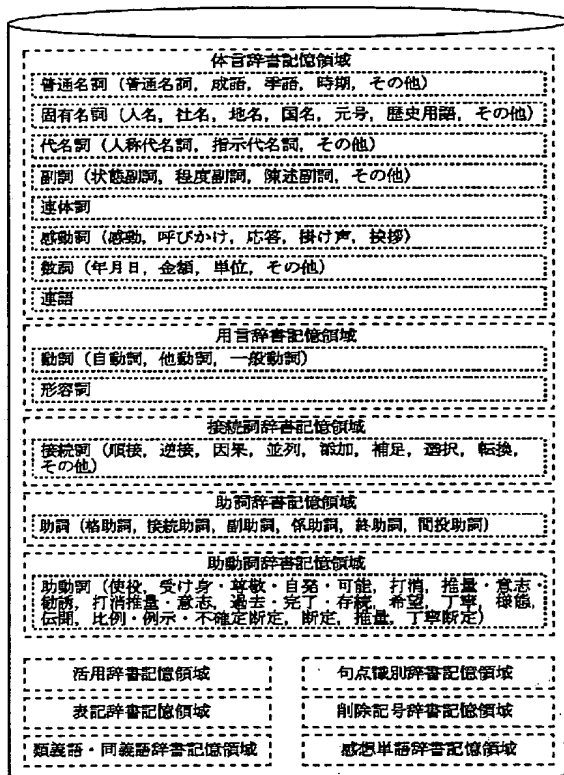
【圖2】



【図5】

[illegible]

【図3】



13

【図9】

	職種		理解		分析			考	表現		応用		
指標値比較	職種A	職種B	理解力	把握力	分析力	洞察力	問題発見力	創造力	表現力	読得力	応用力	行動力	対応力
文章1	×	△	5.1	4.9	5.5	5.2	5.1	4.5	5.2	4.4	5.4	4.3	4.9
文章2	△	×	4.5	4.6	5.1	4.6	4.9	4.0	3.7	3.2	4.5	3.8	4.7
文章3	○	△	5.0	5.5	5.7	5.0	5.2	4.2	5.6	4.5	5.6	4.0	5.3
文章4	×	○	5.9	5.0	6.6	5.8	5.4	4.7	6.1	4.6	6.4	4.5	5.1
文章5	△	○	5.3	5.5	5.6	5.5	5.8	4.7	5.8	4.6	5.5	4.5	5.4
文章6	○	×	4.7	4.3	5.4	5.3	4.5	3.9	5.4	4.3	5.3	3.7	4.3
文章7	○	○	4.8	4.7	5.1	5.5	4.9	4.5	4.9	4.4	5.0	4.3	4.7
文章8	○	△	5.0	5.4	5.3	4.9	5.2	4.4	5.1	4.4	5.2	4.2	5.3
文章9	×	○	4.8	5.0	4.9	4.7	4.9	4.2	4.9	4.3	4.8	4.0	4.9
.
.
.

【図4】

対象	総文字数	計算文字数	総頻度数	用言数	出現単語数	センテンス数	キーワード数	感想単語数	1頻度単語数
文章1	1034	427.1	259	38	186	44	7	5	141
文章2	1625	671.1	394	57	250	56	7	8	172
文章3	1603	664.8	450	45	295	46	16	8	211
文章4	1111	442.5	284	54	218	34	4	5	167
文章5	857	337.8	231	29	155	33	5	5	116
文章6	1439	593.4	380	50	270	49	5	6	208
文章7	1854	733.7	467	61	306	42	10	9	227
文章8	1276	518.2	344	35	214	30	12	6	154
文章9	1801	738.0	479	47	300	37	18	11	222
.
.
.
.

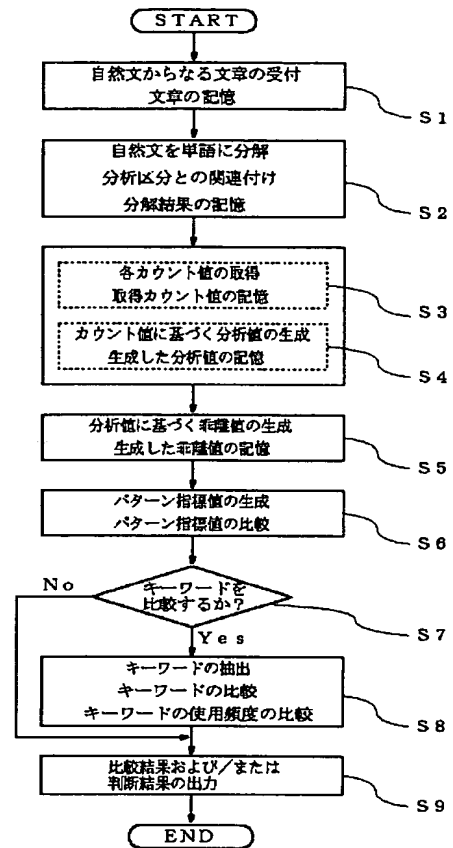
【図6】

7'ログ	第1			第2			第3		第4			第5				
乖離値	体言率	用言率	付属語率	中心キー ワード比率	単語比率 中心補完率	単語比率 中心1頻度	総頻度比率	出現単語 比率	センテンス 用言率	センテンス 頻度平均値	センテンス 比率	第1単語 頻度比率	第2単語 頻度比率	第3単語 頻度比率	重複率 同頻度数	単語回帰値
文章1	5.4	6.5	5.3	4.6	6.8	4.7	5.3	5.3
文章2	3.3	6.2	5.0	3.8	7.5	2.9	5.0	3.8
文章3	6.1	4.8	6.4	5.9	7.5	4.4	6.4	5.4
文章4	6.1	9.1	5.9	3.9	7.9	5.9	5.9	6.6
文章5	5.9	6.2	6.5	4.2	7.0	5.1	6.5	5.8
文章6	5.6	6.1	5.3	3.8	6.9	5.0	5.3	5.7
文章7	4.9	6.0	5.8	3.9	6.8	4.3	5.8	4.8
文章8	5.4	4.8	6.2	5.8	6.9	4.0	6.2	4.8
文章9	5.1	4.5	6.0	5.4	5.9	4.0	6.0	4.6
.
.
.

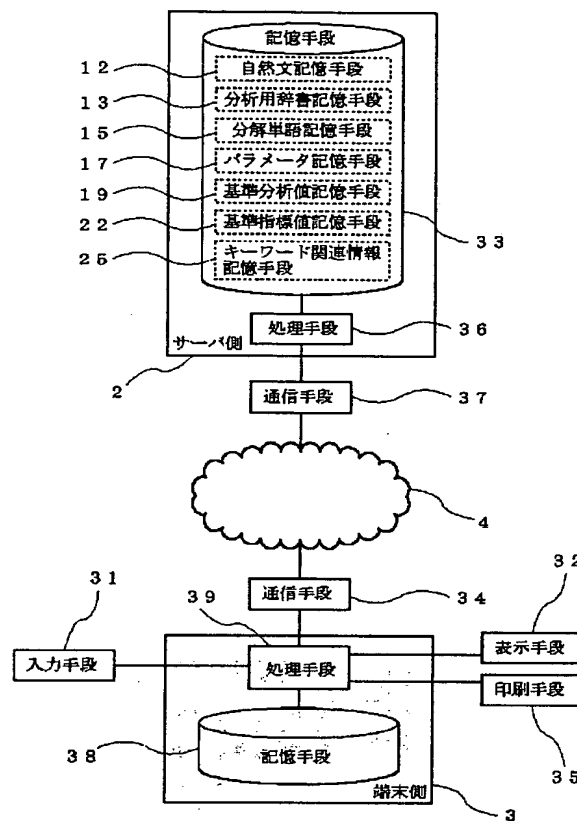
【図7】

7'ログ	第1			第2			第3		第4			第5				
乖離値 対応値	体言率	用言率	付属語率	中心キー ワード比率	単語比率 中心補完率	単語比率 中心1頻度	総頻度比率	出現単語 比率	センテンス 用言率	センテンス 頻度平均値	センテンス 比率	第1単語 頻度比率	第2単語 頻度比率	第3単語 頻度比率	重複率 同頻度数	単語回帰値
文章1	0	+2	-1	0	+2	+1	0	0
文章2	0	+2	+2	0	+2	-1	0	-2
文章3	0	-2	+1	0	+2	-2	0	-1
文章4	0	2	-1	0	+2	+2	0	+1
文章5	0	+1	+1	0	+2	+1	0	-1
文章6	0	+1	-1	0	+2	+2	0	+1
文章7	0	+2	+1	0	+2	+1	0	-1
文章8	0	-1	+1	0	+2	+2	0	-2
文章9	0	-1	+1	0	+1	-2	0	-2
.
.
.

【図8】



【図10】



フロントページの続き

(72)発明者 萬徳 千賀子
東京都渋谷区代々木一丁目10番1号 有限
会社シードウィン内

(72)発明者 永嶋 絃子
東京都渋谷区代々木一丁目10番1号 有限
会社シードウィン内
Fターム(参考) 5B009 QA03